WAVEFORM ANALYSIS

A GUIDE TO THE INTERPRETATION OF PERIODIC WAVES, INCLUDING VIBRATION RECORDS

BY

R. G. MANLEY

B.Sc. (Hons., London)

VIBRATION ENGINEER, SILENTBLOC LIMITED

FORMERLY OF THE VIBRATION DEPARTMENT, DE HAVILLAND AIRCRAFT CO. LTD.

CHAPMAN & HALL LTD.

LONDON

1945

First published 1945

BOOK PRODUCTION WAR ECONOMY STANDARD

THIS BOOK IS PRODUCED IN COMPLETE CONFORMITY WITH THE AUTHORIZED ECONOMY STANDARDS

CATALOGUE No. 289/4

PRINTED IN GREAT BRITAIN
BY THE ABERDEEN UNIVERSITY PRESS LIMITED
ABERDEEN, SCOTLAND
BOUND BY G. & J. KITCAT LIMITED, LONDON
Flexiback Binding

PREFACE

ALTHOUGH there is a tremendous and ever-growing mass of literature concerning those branches of science and engineering which involve the study of periodic variations—mechanical vibration engineering, electrical and radio research and acoustics, for example—experience has shown that there is a lack of practical information on the interpretation of recorded waveforms and allied subjects.

This book has been written in an attempt to make good the deficit. As the sub-title indicates, it is in the nature of a guide, rather than a comprehensive treatise; yet an endeavour has been made to include sufficient details of the basic theory and the methods of analysis to enable the reader to acquire the essential groundwork of knowledge concerning the properties of complex waveforms and to pursue the study of any particular method of analysis to a moderately advanced stage.

Apart from this general intention, the special object in writing the book has been to present in detail the envelope method of analysis which enables recorded waveforms (such as vibration records) to be analysed into their principal constituents without having recourse to mechanical contrivances or to the more exact but cumbersome methods usually known as "Fourier analysis" or "Harmonic analysis". Indeed, the title was deliberately chosen in preference to the more familiar "Harmonic analysis" in order to emphasise the fact that particular reference is made, throughout the text, to the needs of scientists and engineers who may require to interpret by analysis many hundreds or even thousands of recorded waveforms.

The envelope method of analysis was brought to its present state of development largely as a direct result of such a definite need. While in the Vibration Department of the de Havilland Aircraft Company, the author was for some time engaged in the routine analysis of many hundreds of strain-gauge records, taken in an extensive research programme. It was required to estimate, with as great a degree of accuracy as circumstances permitted, the amplitudes and frequencies of the principal constituents of the waves. At the commencement of the work, very little was known of any practicable processes for the analysis. The great number of the

records,' and several special difficulties to which references are made in the text, precluded the possibility of utilising the normal numerical method. Some technique of analysis by inspection was required, and was in due course developed. Even the relatively simple case of beating between two vibrations of nearly equal frequencies caused trouble at the start; the rule for determining the frequency of the minor component, which is not apparent in the wave, had to be discovered. It is well-known that this frequency is either the sum or the difference of the frequency of the major component and the beat frequency; the difficulty lies in determining whether the sum or the difference is to be taken. Although it is hardly credible that no one should have found the rule before, no reference to it has been encountered in the standard literature.

There is no easy road to proficiency in the art of waveform analysis. Experience in the training of young analysts has shown that a thorough appreciation of the properties of complex waves is best obtained by way of a preliminary study of the process of synthesis, and for this reason the first chapter of the book deals with the basic properties of sine waves and the manner in which they combine to form two- and three-component waveforms.

Some suggestions for study, and notes on the plan of the book, are included in the Introduction. With reference to the other methods of analysis, discussed in the later chapters, it is hoped that the periodic existence functions (of which an outline is included in Chapter VI) may be of use in the more mathematical applications of the theory. Attention may be drawn to the unavoidable error in numerical analysis using equally-spaced ordinates (Chapter VII): although this error was first described in print in 1921, and was treated in some detail by Eagle (reference B in the Bibliography) in 1925, none of the modern standard texts on mechanical vibrations or engineering mathematics appears to make any reference to it. Numerical analysis with unequally-spaced ordinates has not been treated, but references to relevant papers will be found in the Bibliography. The last of the five appendices includes a table of the values of the first sixteen sine and cosine harmonics at close intervals, which should serve to eliminate much of the labour of checking an analysis by summing the indicated components and comparing the sum with the original.

The author's thanks are due to the de Havilland Aircraft Com-

pany for permission to publish the results of researches undertaken on its behalf, and to reproduce some specimen records; to Dr. Ker Wilson, to whose persistent encouragement and helpful suggestions the book owes much; to the Director of the Science Museum for permission to use a machine at the museum for the construction of some of the examples of complex waveforms; to the Clarendon Press, and the Editor of "Engineering", for permission to reprint extracts which are acknowledged in greater detail in the text; and to Dr. M. K. B. Richards, but for whose help the numerical work involved would never have been accomplished.

Finally, it is a pleasure to record an appreciation of the never-failing courtesy of the publishers, and the co-operation of the printers, which have considerably eased the task of bringing the book to publication in somewhat trying circumstances.

R. G. Manley.

Mill Hill, September, 1944.

CONTENTS

PLATES.

INTRODUCTION

Specialisation.

At a not very remote period in the past, a university education in Natural Philosophy, together with a small amount of private reading, enabled a man to claim fairly that he knew the whole of science, so far as it was at that time revealed. In contrast, the present-day study of science is so extensive and intensive that no one can hope to acquire a thorough knowledge of more than a small portion of one of the sciences. Specialisation is forced upon the scientific worker who desires to contribute useful original work, for the successful and economical achievement of which it is essential to be well-informed of contemporary progress by other investigators in the same field of study. This unavoidable trend towards learning "more and more about less and less" necessarily involves some considerable dependence upon the results of research and development in other subjects; and perhaps the greatest disadvantage of the situation is that one has to accept without question these results, as time alone prevents a thorough *ab initio* investigation of principles and methods from being made.

While the foregoing remarks are broadly true of science in general, there is a conspicuous exception to the rule in the case of mathematics. Mathematics is, in its utilitarian aspect, the handmaid of all the sciences, in that it provides a set of processes for solving problems posed in their most general terms. Everyday life is permeated with the use of figures, and a mathematical undercurrent is observable in any study of physical phenomena. It is not surprising, therefore, that workers in very different fields make use of identical or similar mathematical processes, in the solution of problems which have a common mathematical nature, although they may differ widely in physical significance.

Fourier series.

A relevant example is the case of the study of periodic variations. It is impossible to give an exhaustive list of physical phenomena which give rise to periodic variations; some of the most important are mechanical vibrations, alternating current electricity, acoustics and tidal motions. All these involve variations which are repetitive after successive equal intervals of time, and to which the general name "periodic variations" is given; the analysis and interpretation of these variations depends upon a theorem which is inseparably linked with the name of Fourier.

To anyone conversant with the powers and peculiarities of mathematics it will not seem strange that, in fact, Fourier's discovery was made in connection with the study of *non-periodic* phenomena in the conduction of heat.

Restricting present attention to periodic phenomena, the practical significance of Fourier's Theorem is that any periodic variation which satisfies certain conditions of continuity (which are sufficiently favourable as to include practically all physically possible cases) can be regarded as the sum of a number of simpler periodic variations of a special kind. These are in the form of sine waves, whose periods form a harmonic series in the mathematical sense, i.e. their reciprocals form an arithmetic series or progression. According to the form of the original variation, the constants associated with these sine wave components, commonly termed *harmonics*, are uniquely determined by the process of *Harmonic Analysis*.

Provided the validity of this representation is accepted, the determination of the constants is almost absurdly simple so long as the variation can be stated in the form of a mathematical function, or set of such functions, and certain integrations can be performed. The well-known formulæ for harmonic analysis give the values of the constants explicitly, and the only problem that can arise is that of evaluating the integrals. This mathematical variety of harmonic analysis is used very widely in physics and engineering science (see, for example, reference 1 in the Bibliography).

Numerical analysis.

In many cases, however, the variation to be analysed is not given in the form of a mathematical function. Indeed, it may be desired to find a suitable function which may represent, with a sufficient degree of accuracy, a given set of numerical values for a periodic variation, the values usually having been obtained experimentally. For this purpose the mathematical formulæ are modified, approximations to the integrals being found by direct summation. This numerical variety of analysis is well-suited to tabulation; Runge (reference 2) appears to have been the first to give convenient tabular schemes which have been at all widely employed.

In both the pure mathematical, and the modified numerical, forms of analysis the aim is to determine as accurately as possible the harmonic contents of the variation, for a variety of purposes. Engineering examples in the realm of vibration study illustrate both procedures: the accelerations of the pistons in an internal combustion engine, when the crankshaft is rotating uniformly, are analysed by the straightforward mathematical process in order to

assess the degree of unbalance of the corresponding inertia forces; and the gas pressure in the cylinders of the engine is subjected to numerical analysis, by means of ordinates measured off an indicator diagram, to determine the degree of unbalance of gas forces, particularly those components which initiate and maintain the insidious torsional vibrations of the crankshaft system.

Envelope method.

There remains, however, an extensive class of periodic variations which, for various reasons, cannot readily be analysed by either of the methods just described. The outstanding examples of this class are recorded waveforms depicting mechanical vibrations. In a comprehensive vibration test there may be many hundreds of such records, which are obtained from vibrographs, torsiographs, or strain gauges by means of auxiliary electronic equipment and a recording oscillograph. For the determination of the harmonic contents of these waveforms a third type of analysis has been evolved, and it is this type which it is the principal aim of this book to describe and discuss. The procedure is known generally as the "envelope method."

The distinguishing features of the envelope method of analysis are that it enables the main properties (frequency, amplitude and phase) of the principal components of the waveform to be determined very speedily, and fairly accurately, by means of auxiliary constructed lines termed "envelopes." The great majority of recorded waveforms obtained in vibration study, and in other investigations, contain only two or three principal components, and the envelope method is competent to handle most varieties of such waveforms. In special cases, waveforms containing four or five principal components can be analysed, but the accuracy of the results deteriorates as the complexity of the waveform increases. A further special feature of the method is the fact that it can be applied in cases where the cycle of the waveform is very long, or even when no true cycles exist owing to non-periodic fluctuations or the disturbing influence of a non-harmonic component.

Suppose, for example, that the vibrations of a diesel engine are recorded. The engine runs at 800 R.P.M., and it has a four-stroke cycle, so that the frequency of its vibration pattern is 400 C.P.M. If it is further supposed that the recording apparatus is electronic in nature, and is operated from A.C. mains at 3000 C.P.M., there is the possibility of a "mains hum" at this frequency being superimposed upon the record, one cycle of which must contain whole numbers of cycles both of the engine vibration and of the hum.

In this case, an interval of 1/200 minute comprises simultaneously two cycles of the engine vibration and fifteen of the hum. If, however, the speed of the engine fluctuates to 820 R.P.M., so that the frequency of the vibration pattern is 410 C.P.M., the cycle of the resultant recorded waveform extends over a period of six seconds, equivalent to 41 cycles of engine vibration and 300 cycles of hum; the harmonic of lowest order is the 41st. Again, as the engine speed is likely to fluctuate slightly, the waveform will probably show no definite cycles at all.

Waveforms of this type cannot be analysed by the numerical method, or by means of the various types of mechanical aids. The envelope method can, however, be applied successfully to such waveforms, and is used exclusively in one large industrial establishment which handles thousands of strain-gauge records every year.

Other methods.

Various machines have been proposed for the purpose of relieving the drudgery of waveform analysis, and for reducing the time taken in the analysis of large numbers of records.

An alternative method of calculating the Fourier coefficient in a certain class of waveforms is known as the "method of superposition." The cycle of the waveform is divided in different ways into sets of sub-cycles, addition or subtraction of which separates out different sets of harmonics. Thus, if the cycle is divided into two half-cycles, addition of the two parts yields a variation which is twice the sum of the even harmonics, while subtraction of one from the other yields a variation which is twice the sum of the odd harmonics.

Suggestions for study, and plan of book.

For whatever reason the study of waveform analysis is made, the best plan is to concentrate first on the opposite process of the synthesis of a complex waveform from its constituents. The first chapter of this book accordingly describes the manner in which sine waves combine; and, so valuable is a thorough grasp of the mechanism of synthesis, that this chapter may be regarded as forming the most important part of the text.

The reader is strongly urged to build-up many examples of complex waveforms from their components by numerical synthesis. For this purpose, the use of the normal trigonometrical tables is highly inconvenient, and in Appendix V has been included a table giving the values of the first sixteen sine and cosine harmonics at intervals of 1/96 of the fundamental cycle. The use of this table greatly expedites the synthesis.

The second chapter, in which are set out the main properties of harmonic series, completes the preliminary study and prepares the way for the detailed study of the various methods of analysis.

Chapter III describes the basic analysis of recorded waveforms, and includes detailed instructions for the determination of frequencies, amplitudes and phase-angles of sine-waves. Thereafter, the various methods of analysis are described and discussed in turn. The reader will naturally select those chapters which are particularly relevant to the method or methods he wishes to use, but it may be as well to point out that there are many practical hints included in the chapter on the envelope method (Chapter IV) which can sometimes be utilised with advantage in analysis by other methods.

The main body of the text is completed by two chapters, dealing with practical requirements for waveforms, and with Lissajou figures. In the former, such subjects as clarity of records, recording media, linearity of response, length of records and time-reference markings are discussed, the aim being to simplify the analysis. The work of the analyst can be considerably facilitated by due attention to these practical matters, and although in small research establishments the analyst is likely also to be the person responsible for design of equipment and technique of testing, in many cases the present-day trend towards specialisation, together with a marked expansion of research departments, has resulted in a division of these activities between different individuals, between whom there may not be any very close co-operation.

The final chapter, on Lissajou figures, has been included to render the text more complete; the figures, while being completely different in nature from the usual recorded waveforms, nevertheless are examples of periodic waveforms.

Certain subsidiary matters and general considerations have been included in the appendices, including the tables of the sine and cosine harmonics for use in numerical synthesis.

Terms and phrases peculiar to the study of periodic variations have been defined in the text, but for convenience these definitions, and also those of certain other terms of frequent occurrence, have been collected together in the glossary on pages 262 to 265.

CHAPTER I

SINE-WAVES IN COMBINATION

1. Introductory.

The practical result of Fourier's Theorem (see Chapter VI) is that any periodic variation whatever, i.e. any function which is exactly repetitive after successive equal intervals of time, or of whatever the basic variable may be, can be expressed as the sum of a number of sinusoidal variations. In conjunction with the sinusoidal character of the phenomena associated with mechanical vibration, alternating-current electricity, acoustics and optics, this fact concentrates interest and study on the behaviour of the sine function, and particularly on its graphical representation as a sine-wave.

Wave-form analysis consists of determining the appropriate sine-waves whose sum represents the given wave-form. It is properly termed "analysis," since it is a process of breaking-down, and although in the mathematical treatment of the problem no knowledge is required of the inverse process of synthesis—i.e. building-up—yet in the practical treatment with which part of the present work is concerned a study of the results obtained by adding together different sine functions is very helpful in the development of an analysis technique.

This chapter is therefore concerned with the general properties of sine-waves, and with the waveforms that can be produced by adding two different sine-waves together; at the end of the chapter a cursory examination is made of three-component waveforms. In subsequent chapters these results are extended and used to derive a speedy and accurate method of analysis for waveforms involving two, three or four principal components.

Particular interest is focussed on waveforms in which the components are of the harmonic type, i e. their frequencies are all integral multiples of a fundamental frequency, since such series of components occur universally in practical problems. It may frequently happen, however, that harmonics of two different fundamental frequencies are present in the same waveform, and the reader should observe that except where an exactly integral frequency ratio is expressly stipulated in this chapter the conclusions are valid for the addition of two or three sine-waves whose frequency ratios cannot be expressed as a simple ratio of integers.

2. Sine function.

For the purposes of the study of wave-forms the sine function is best defined geometrically from the properties of rotating vectors. Let the point P move on a circle, whose centre is O (Fig. 1), and let rectangular axes OX, OY be chosen in the plane of the circle. As the point P moves round the circle the direction of OP varies, and OP is said to be a rotating vector, a vector being any quantity which has direction as well as magnitude. The length l of OP is considered to be positive whatever position P may lie in. Let N, Q be the feet of the perpendiculars drawn from P to OX, OY respectively; and let the normal rule of signs be observed, i.e. that ON is taken to be positive or negative as N lies to the right or left of OY, and OQ is taken to be positive or negative as Q lies above or below OX. The sine and cosine functions are defined thus:

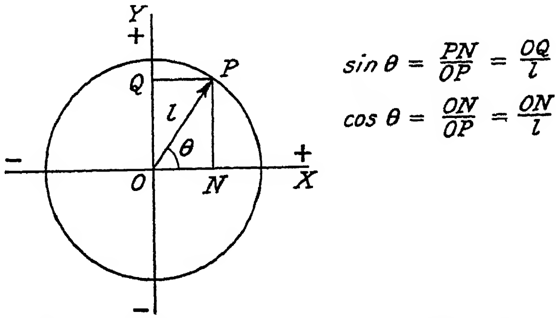


Fig. 1.—Definition of sine and cosine functions.

$$\left.\begin{aligned} \sin\theta &= \frac{\mathrm{NP}}{\mathrm{OP}} = \frac{\mathrm{OQ}}{l} \\ \cos\theta &= \frac{\mathrm{ON}}{\mathrm{OP}} = \frac{\mathrm{ON}}{l} \end{aligned}\right\} \quad . \quad . \quad . \quad (2.1)$$

where θ is the angle XOP measured in the counter-clockwise sense from OX.

Note.—Both the radian measure and the degree measure are useful on different occasions in this work. The mathematician prefers the radian measure, as it is less arbitrary than the degree measure, but to engineers and other practical investigators a right-angle is more clearly indicated if it is called "90 degrees" than if the term "$\pi/2$ radians" is employed. In this book both measures are used, preference being given to the degree measure wherever its use is possible, for the purpose of simplicity.

If the length l of the vector is made unity, the sine and cosine functions are then seen to be defined as the projections of this unit vector upon the two axes of co-ordinates.

From geometrical considerations the well-known fundamental identities are obtained:

$$\sin^2 \theta + \cos^2 \theta = 1 \qquad . \quad . \quad . \quad . \quad (2.2)$$

and
$$\sin(\theta + 90^\circ) = \cos \theta \qquad . \quad . \quad . \quad . \quad (2.3)$$

The following general formula is proved in any textbook of elementary trigonometry :

$$\sin(\theta + \phi) = \sin \theta \cos \phi + \cos \theta \sin \phi \qquad . \quad . \quad (2.4)$$

and this result, together with the special values, listed in Table I, of the sine and cosine functions for angles which are multiples of 90°, yields some useful relations :—

$$\left.\begin{aligned}
\begin{pmatrix}\sin\\ \cos\end{pmatrix}(\theta + 90^\circ) &= \begin{pmatrix}\cos\\ -\sin\end{pmatrix}\theta^{*}\\
\begin{pmatrix}\sin\\ \cos\end{pmatrix}(\theta + 180^\circ) &= -\begin{pmatrix}\sin\\ \cos\end{pmatrix}\theta\\
\begin{pmatrix}\sin\\ \cos\end{pmatrix}(\theta + 270^\circ) &= \begin{pmatrix}-\cos\\ \sin\end{pmatrix}\theta\\
\begin{pmatrix}\sin\\ \cos\end{pmatrix}(\theta + 360^\circ) &= \begin{pmatrix}\sin\\ \cos\end{pmatrix}\theta\\
\begin{pmatrix}\sin\\ \cos\end{pmatrix}(-\theta) &= \begin{pmatrix}-\sin\\ \cos\end{pmatrix}\theta
\end{aligned}\right\} \qquad . \quad . \quad (2.5)$$

TABLE I

Values of sine and cosine functions of angles which are multiples of 90°

$\theta =$ . . .	0° 360° 720° etc.	90° 450° 810° etc.	180° 540° 900° etc.	270° 630° 990° etc.
Sin θ . .	0	1	0	−1
Cos θ . .	1	0	−1	0

A collection of useful trigonometrical formulæ, and a table of sines and cosines of multiples of 15°, are given in Appendix V, p. 256.

3. Generation of sine-wave ; definitions.

Period, cycle. Let the point P in Fig. 1 move so that the angular velocity of OP is ω radians per second in the counter-clockwise sense, and let the time t seconds be measured from the instant when

* The notation here employed is useful to condense two similar equations into one statement. Thus the first equation in (2·5) represents the two relations :

$$\sin(\theta + 90^\circ) = \cos \theta, \ \cos(\theta + 90^\circ) = -\sin \theta.$$

$\theta = \phi$. Then $\theta = \omega t + \phi$, and a graph of $l \sin \theta$ plotted against t has the form shown in Fig. 2. The graph is repetitive after successive intervals of time $2\pi/\omega$ seconds, since this is the time taken for the point P to complete one circuit and regain its original position; this interval is termed the *period* of the function, which is said to be *periodic*. The term "period" applies only to the basic variable t, the corresponding part of the graph being known as a *cycle*.

Descriptive terms. The terms *crest* and *trough*, used in connection with waveforms, are self-explanatory. Both crests and troughs are sometimes referred to as *peaks*, however, probably because in the interpretation of many recorded waveforms it is immaterial which is considered the "right way up," and a crest becomes a trough when the wave is inverted.

The axis ΘX, being the axis of the basic variable t, is commonly

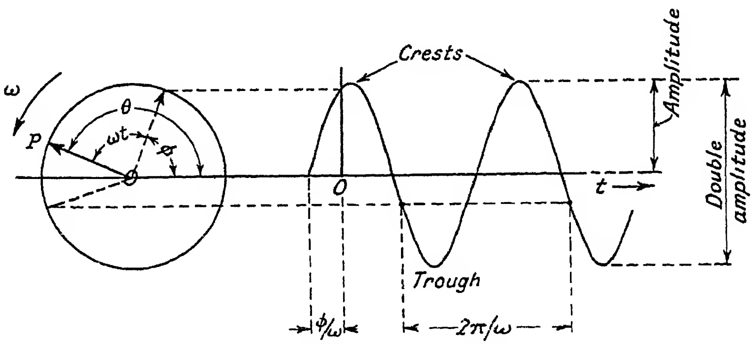


Fig. 2.—Generation of sine-wave as projection of rotating vector.

termed the *centre-line*. This name is also applied in the case of more complicated waveforms, when it refers to a line parallel to the axis of the basic variable and indicating the average or mean height of the wave; thus in a waveform representing the function $3 + 2 . \sin x$ the centre-line is parallel to the x-axis and at a height above it which represents 3 on the scale to which the waveform is drawn, since the average value of the sine function over a cycle is zero.

Amplitude. Half the total height of the wave (i.e. the length of OP) is termed the *amplitude*, and the total height is frequently termed the "double amplitude."

Phase-angle. The angle ϕ is termed the *phase-angle*, and the sole effect of a change in the phase-angle is to shift the wave along the time-axis. Some typical examples of the phase-effect are shown in Fig. 3.

1 *

Although in vibration study it is sometimes convenient to allow amplitudes to take negative values, by restricting the phase-angles to lie within the range 0-180° (see reference 1 in the Bibliography), in waveform analysis it is better to regard all amplitudes as positive, and to allow phase-angles to take any value.

$$2 \,.\, \sin(\omega t + 257^\circ) = (-2)\sin(\omega t + 77^\circ)$$
$$8 \,.\, \sin(\omega t + 270^\circ) = (-8)\sin(\omega t + 90^\circ)$$
$$= (-8)\cos \omega t.$$

Two waves having the same period have a constant *phase-difference*, which is the difference between the two phase-angles

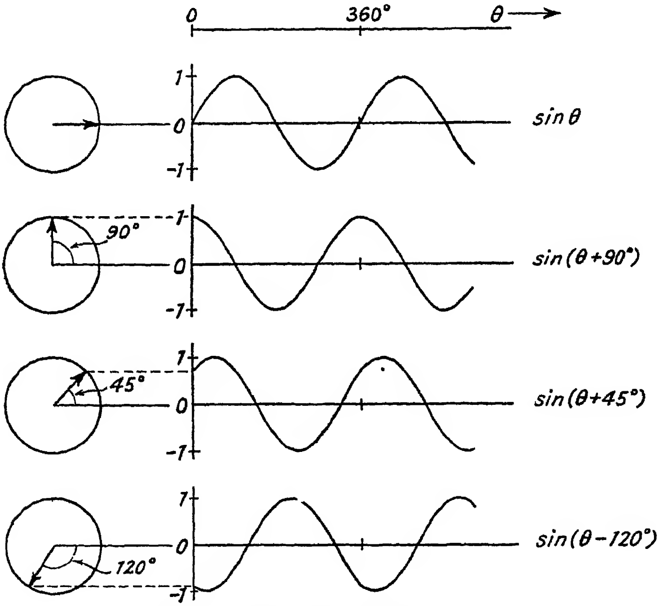


FIG. 3.—Examples of phase-shift.

(i.e. the angle between the two rotating vectors which are associated with the waves). Thus the phase-difference between the waves $\sin(\omega t + 5^\circ)$ and $\sin(\omega t - 47^\circ)$ is 52°, since at the instant $t = 0$ the vector associated with the former wave includes a positive angle of 5° with the t-axis, while the vector associated with the latter includes a negative angle of 47°.

Two waves whose phase-difference is 180° or π radians are said to be *anti-phased.* Except for amplitudes, one wave is then the mirror-image, in the centre-line, of the other, as is seen from the equation $\sin(\theta + 180^\circ) = -\sin\theta$. The vectors which generate the two waves are in opposite senses, as their angular separation

is 180°, and a crest in one wave occurs simultaneously with a trough in the other. If the phase-difference between two waves is zero, they are said to be *in-phase*. The vectors which generate the two waves are then parallel and in the same sense, and crests occur simultaneously in the two waves; in fact, one wave is simply a magnified version of the other, the measurements parallel to the centre-line being unaffected.

The term "phase-difference" is only of significance when applied to two waves of equal period, since if the periods of two waves are not the same there is no fixed angle between the rotating vectors. Nevertheless, it is sometimes useful to distinguish between different types of phase-relation in the case of two waves of unequal periods. No matter what the values of ω_1 and ω_2 may be, the waves $a_1 \sin \omega_1 t$

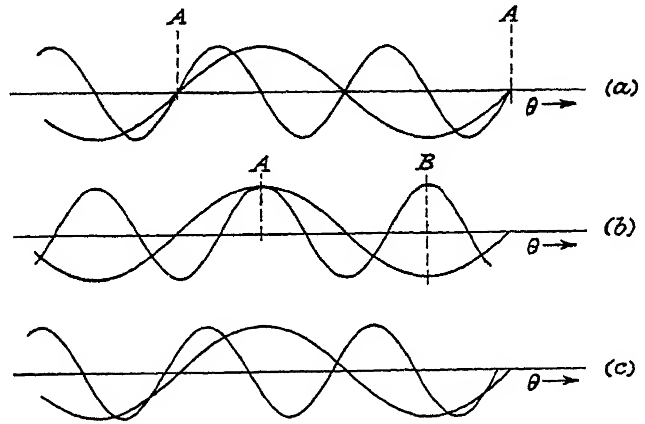


Fig. 4.—Phase relations between waves of different frequencies: (*a*) simultaneous zeros; (*b*) simultaneous peaks; (*c*) neither of these properties.

and $a_2 \sin \omega_2 t$ both have a zero value at the instant $t = 0$ and again at the instants $t = \mathrm{T}$, 2T, 3T, etc., where T is the smallest time-interval which includes an integral number of both the semi-periods π/ω_1 and π/ω_2. Fig. 4*a* shows the two waves $\sin \theta$ and $\sin 2\theta$, both having zero points at A. Similarly, both the waves $a_1 \cos \omega_1 t$ and $a_2 \cos \omega_2 t$ have a maximum value at the instant $t = 0$, and again at the instants $t = 2\mathrm{T}$, 4T, etc., while at the instants $t = \mathrm{T}$, 3T, etc., they are temporarily anti-phased, one having positive maximum value while the other has negative maximum value. Fig. 4*b* shows the two waves $\cos \theta$ and $\cos 2\theta$, A being a point where they are temporarily in-phase, and B being a point where they are temporarily anti-phased. Fig. 4*c* illustrates a case where the phase-relation is of neither of these types, (*a*) or (*b*). It should be noted that the point

A in (*a*), denoting the instant $t = 0$, is not vertically over the similar point in (*b*). This lateral displacement has been introduced deliberately, as it is important to bear in mind the desirability of an arbitrary choice of datum-point. This matter is discussed in Chapter II (p. 50).

Frequency. The reciprocal of the period is termed the *frequency*, and is measured in *cycles per second* (C.P.S.) * or *cycles per minute* (C.P.M.). Strictly, this name and "period" should be used only when the basic variable is a time measurement, and the use of the term "frequency" as a quantitative term is normally so restricted; for convenience, however, "period" is frequently employed for functions whose basic variable is other than a time measurement —for example, it is said that 2π radians is the period of the sine function $\sin\theta$—and the terms "high-frequency" and "low-frequency" are often employed with reference to functions whose periods are, respectively, relatively small or relatively great, no matter what the basic variable may be.

The relation between the frequency and the angular velocity ω of the rotating vector, sometimes called the *phase-velocity*, is given by

$$\left.\begin{aligned}\text{frequency F} &= \omega/2\pi \text{ C.P.S.*}\\ &= 9{\cdot}55\omega \text{ C.P.M.}\end{aligned}\right\} \quad . \quad . \quad (3.1)$$

When frequencies are very great it is inconvenient to state them in C.P.M. or even in C.P.S., owing to the large numbers involved. Such high frequencies are measured in *kilocycles per second* (a kilocycle being one thousand cycles) or *megacycles per second* (a megacycle being one million cycles). The abbreviations used for these units are Kc. and Mc. respectively. Thus:

1 Kc. = 1,000 C.P.S. = 60,000 C.P.M.,

and 1 Mc. = 1,000,000 C.P.S. = 60,000,000 C.P.M.

Note.—Of all these descriptive terms and quantities associated with waveforms, *frequency and amplitude* are the most important, whether the waveform results from oscillatory phenomena in mechanical vibration, electricity, acoustics or optics. In mechanical vibration analysis the two terms recur again and again; the author has given elsewhere (reference 2 in the Bibliography) an account of the practical importance of amplitudes and frequencies in this study. It is sufficient here to state that the development of an accurate and speedy method of determining the amplitudes and frequencies of all the component vibrational variations in a structure is of vital importance in the solution of mechanical vibration problems. *Phase-angles* are generally of less practical importance; nevertheless, in many

* Continental writers sometimes employ the unit "hertz," 1 hertz being 1 C.P.S.

investigations the phase-characteristics of a wave may be important, particularly in the theoretical estimation of the vibrational energy-input in a system; in evaluating this quantity consideration must be given to the phase-difference between the excitation force and the displacement produced by it. (See reference 3 in the Bibliography.)

4. Symmetry, skew-symmetry, and alternance.

The sine-wave exhibits the three properties of *symmetry*, *skew-symmetry*, and *alternance*.

(i) *Symmetry*.

Since $\sin(180° - \theta) = -\sin(-\theta) = \sin\theta$, it follows that the graph of $\sin\theta$ is symmetrical about the values $\theta = 90°$, $270°$, etc.: i.e. if K is any integer,

$$\sin[(2K - 1)\,90° + \alpha] = \sin[(2K - 1)\,90° - \alpha] \quad . \quad (4.1)$$

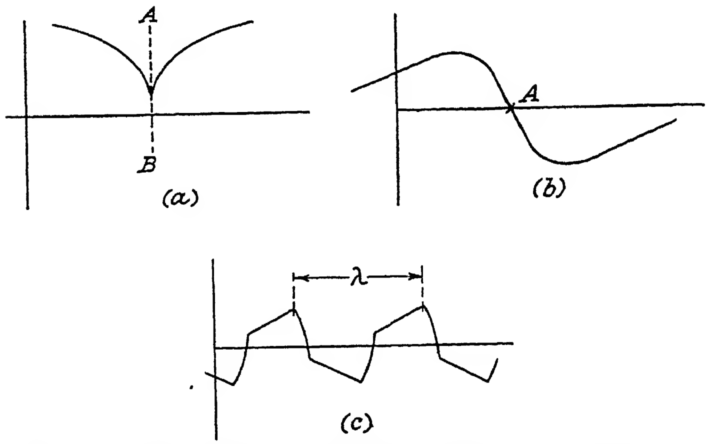


FIG. 5.—(*a*) Symmetry; (*b*) skew-symmetry; (*c*) alternance.

for all values of α. Thus if the paper on which the graph is drawn is folded along any of the lines $\theta = 90°$, $270°$, etc., the two parts of the graph so divided will touch along the entire length of the curve.

To illustrate the property of symmetry more simply, Fig. 5*a* depicts a non-periodic function which is symmetrical about the dotted line AB.

(ii) *Skew-symmetry*.

Since $\sin\theta$ is symmetrical about the value $\theta = 90°$, and $\sin(-\theta) = -\sin\theta$, it follows that the function possesses skew-symmetry about the values $\theta = 0$, $180°$, $360°$, etc.: i.e. if K is any integer,

$$\sin(K\,.\,180° + \alpha) = -\sin(K\,.\,180° - \alpha), \quad . \quad (4.2)$$

for all values of α. The curve of Fig. 5*b* shows part of the graph of a non-periodic function which is skew-symmetrical about the point A.

(iii) *Alternance.*

Since $\sin(\theta + 180°) = -\sin\theta$, the function $\sin\theta$ also possesses the property of alternance with the interval 180° or π radians: i.e. two ordinates separated by a distance representing 180° on the θ axis of the graph are equal in magnitude but opposite in sign. The property of alternance is essentially one possessed only by periodic functions; Fig. 5*c* illustrates such a function, whose period or wavelength is λ, and which is alternant with the interval $\lambda/2$.

These properties are of great use in the process of analysis, as by considerations of the presence or absence of them in a given waveform much information may be gained as to the constituent components. (See Chapter II, p. 60).

5. Addition of two waves of equal periods.

(i) $a \,.\, \cos\theta + b \,.\, \sin\theta$.

The simplest example of the addition of two sine-waves is that wherein the two waves have equal periods and a phase-difference of 90°, i.e. of a quarter-period; such a pair of waves can be expressed quite generally by

$$y_1 = a \,.\, \cos\theta,$$
$$y_2 = b \,.\, \sin\theta,$$

since θ may be chosen arbitrarily. Thus, suppose the two waves are of the form

$$y_1 = a\sin(x + \phi + 90°), \quad y_2 = b\sin(x + \phi).$$

Putting $\theta = x + \phi$ these expressions reduce to $y_1 = a\sin(\theta + 90°)$, $y_2 = b\sin\theta$; and $\sin(\theta + 90°) = \cos\theta$.

Let y be the sum $y_1 + y_2$; then

$$\frac{y}{\sqrt{a^2+b^2}} = \frac{a}{\sqrt{a^2+b^2}}\cos\theta + \frac{b}{\sqrt{a^2+b^2}}\sin\theta.$$

From equation (2.2) it follows that an angle ψ can always be found such that

$$\frac{a}{\sqrt{a^2+b^2}} = \sin\psi, \quad \frac{b}{\sqrt{a^2+b^2}} = \cos\psi,$$

and from equation (2.4) it can be seen that

$$y = (\sin\psi\cos\theta + \cos\psi\sin\theta)\sqrt{a^2+b^2},$$

i.e.
$$\left.\begin{aligned} y_1 + y_2 &= r \,.\, \sin(\theta + \psi) \\ \text{where} \quad r &= \sqrt{a^2+b^2} \\ \text{and} \quad \tan\psi &= a/b \end{aligned}\right\} \qquad (5.1)$$

Thus the wave-form resulting from the addition of sine and cosine waves of equal periods is sinusoidal and has the same period as have the component waves.

In determining ψ from the third equation of (5·1), care must be taken to select the correct value; for *two* possible values are determined by the equation, since

$$\tan(\psi + 180°) = \frac{\sin(\psi + 180°)}{\cos(\psi + 180°)} = \frac{-\sin\psi}{-\cos\psi} = \tan\psi.$$

The appropriate value is easily found by considering the signs of a and b; for $\sin\psi$ is positive if ψ lies between 0° and 180°, negative if ψ lies between 180° and 360°, and $\cos\psi$ is positive if ψ lies between 270° (i.e. $-$ 90°) and 90°, negative if ψ lies between 90° and 270°. Table II gives these signs in a concise form.

TABLE II

Signs of sine and cosine functions in the various quadrants

Quadrant	Range of ψ	Sine	Cosine
1st	0°- 90°	+	+
2nd	90°-180°	+	−
3rd	180°-270°	−	−
4th	270°-360°	−	+

Example.

Suppose the two waves are

$$y_1 = 1.2 \sin(pt - 38°),$$
$$y_2 = 0.5 \sin(pt + 52°).$$

Putting $\theta = pt - 38°$,

$$y = y_1 + y_2 = 1.2 \sin\theta + 0{\cdot}5 \cos\theta.$$

Thus in equations (5.1),

$$a = 0.5, \quad b = 1.2,$$

and $r = 1.3, \quad \tan\psi = 0{\cdot}4167,$

whence $\psi = 23°.$

Thus $y = 1.3 \sin(\theta + 23°) = 1.3 \sin(pt - 15°).$

Fig. 6 shows the graphs of y_1, y_2, and $y = y_1 + y_2$.

The process of addition may also be performed on the vectors which represent the sine-waves. Let OP_1 in the diagram (Fig. 7*a*) be the vector associated with the wave $b \,.\, \sin\theta$, so that the length of OP_1 is b and the angle XOP_1 is θ; then OP_2 being the vector associated with the wave $a \,.\, \cos\theta = a \,.\, \sin(\theta + 90°)$, the length of OP_2 is a,

and the angle XOP_2 is $\theta + 90°$. Thus the angle P_1OP_2 is a right-angle. The resultant vector OR, which is the sum of the vectors OP_1 and OP_2, is found in the same manner as if OP_1 and OP_2 represent velocities, accelerations, or forces—i.e. by the "triangle of vectors" or "parallelogram of vectors" rule. Thus if the length of OR is r, and the angle P_1OR is ψ,

$$r = \sqrt{a^2 + b^2}, \quad \text{and} \quad \tan \psi = a/b.$$

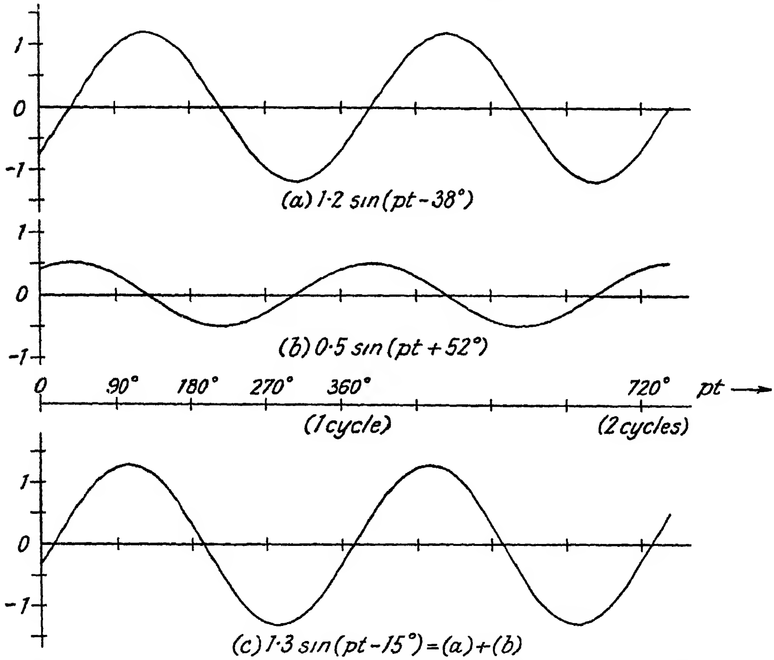


FIG. 6.—Addition of two waves with a quarter-period phase difference.

It is evident that this procedure gives the same result as does the method given above. The vector addition corresponding to Fig. 6 is illustrated in Fig. 7*b*. The vectors shown are those at the instant $t = 0$, and thus occupy angular positions at $-38°$, $52°$, and $-15°$.

(ii) $a \,.\, sin\,(\theta + \phi_a) + b \,.\, sin\,(\theta + \phi_b)$.

The general case of the addition of two sine-waves of equal periods, i.e. where the two waves are of the form :

$$y_1 = a \,.\, \sin\,(\theta + \phi_a),$$
$$y_2 = b \,.\, \sin\,(\theta + \phi_b),$$

is easily broken down into the simpler form (i) : for

$$y = y_1 + y_2 = (a \,.\, \cos \phi_a + b \,.\, \cos \phi_b) \sin \theta$$
$$+ (a \,.\, \sin \phi_a + b \,.\, \sin \phi_b) \cos \theta,$$

by reason of equation (2.4), and if $y = r \,.\, \sin(\theta + \psi)$,

$$r^2 = (a \,.\, \cos\phi_a + b \,.\, \cos\phi_b)^2 + (a \,.\, \sin\phi_a + b \,.\, \sin\phi_b)^2$$
$$= a^2 + b^2 + 2ab\,(\cos\phi_a \cos\phi_b + \sin\phi_a \sin\phi_b),$$

since $\sin^2\phi_a + \cos^2\phi_a = \sin^2\phi_b + \cos^2\phi_b = 1$. Now from equations (2.4) and (2.5) it can easily be shown that

$$\cos\phi_a \cos\phi_b + \sin\phi_a \sin\phi_b = \cos(\phi_b - \phi_a),$$

and hence

$$r = \sqrt{a^2 + b^2 + 2ab \,.\, \cos(\phi_b - \phi_a)}. \qquad (5.2)$$

The phase-angle ψ is given by

$$\tan\psi = \frac{a \,.\, \sin\phi_a + b \,.\, \sin\phi_b}{a \,.\, \cos\phi_a + b \,.\, \cos\phi_b}. \qquad (5.3)$$

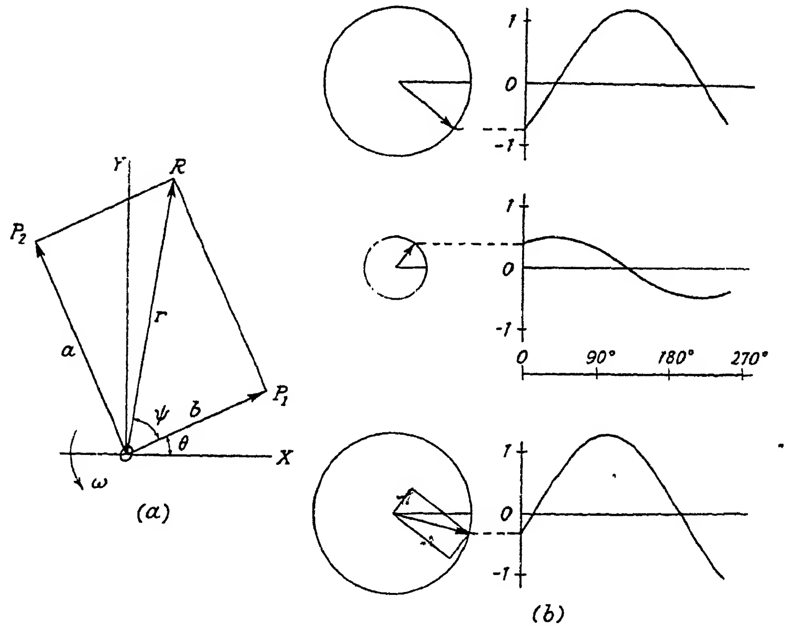


FIG. 7.—Vector representation of addition shown in Fig 6.

The sum $y = y_1 + y_2 = r \,.\, \sin(\theta + \psi)$ where r and ψ are given by equations (5.2) and (5.3).

The method of vector addition is applicable also to this more general case. Let OP_1, OP_2 in Fig. 8*a* represent the vectors associated with y_1 and y_2, so that the angle P_1OP_2 is equal to the phase-difference $\phi_b - \phi_a$. Then by trigonometry the length r of the resultant vector OR is given by

$$r^2 = a^2 + b^2 - 2ab \,.\, \cos[180° - (\phi_b - \phi_a)]$$
$$= a^2 + b^2 + 2ab \,.\, \cos(\phi_b - \phi_a).$$

This result agrees with equation (5.2). The phase-angle XOR $= \psi$ is given by

$$\tan \psi = \frac{\text{RQ}}{\text{OQ}} = \frac{\text{RM} + \text{MQ}}{\text{ON} + \text{NQ}}$$

$$= \frac{b \,.\, \sin \phi_b + a \,.\, \sin \phi_a}{a \,.\, \cos \phi_a + b \,.\, \cos \phi_b},$$

which agrees with equation (5.3).

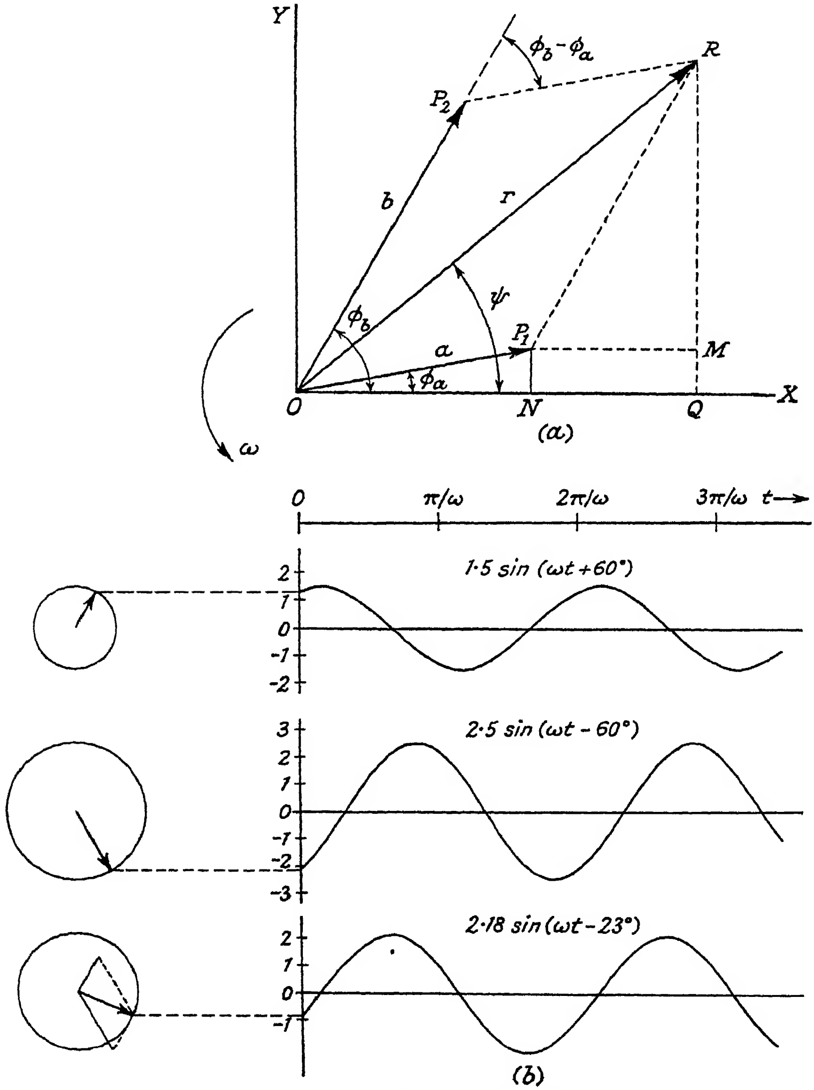


FIG. 8.—Addition of two sine waves of equal periods—general case.

Example.

Suppose that $y_1 = 1{\cdot}5 \sin(\omega t + 60°)$

$y_2 = 2{\cdot}5 \sin(\omega t + 300°)$.

Then $r = \sqrt{2{\cdot}25 + 6{\cdot}25 + 7{\cdot}5 \cos 240°}$,

$= \sqrt{4{\cdot}75} = 2{\cdot}179$,

and
$$\tan \psi = \frac{1{\cdot}5 \sin 60° + 2{\cdot}5 \sin 300°}{1{\cdot}5 \cos 60° + 2{\cdot}5 \cos 300°}$$
$$= \frac{1{\cdot}5 - 2{\cdot}5}{1{\cdot}5 + 2{\cdot}5} \tan 60° = -0{\cdot}433.$$

Thus, since $\sin \psi$ is negative and $\cos \psi$ is positive, $\psi = 337°$.

Fig. 8*b* illustrates the addition.

6. Addition of two waves of different periods: beating.

It has been shown in Section 5 that the sum of two sinusoidal waveforms having equal periods is another sinusoidal wave of the

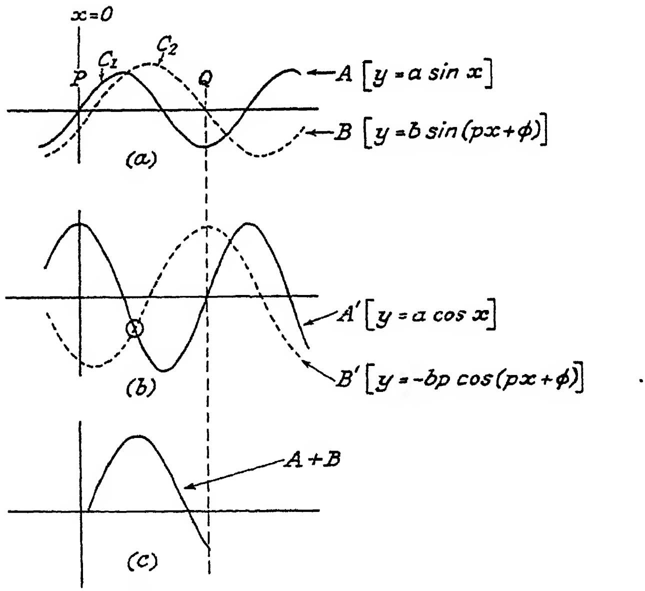


FIG. 9.—Merging of peaks in addition of two waves.

same period. When two waves whose periods are not equal are added together the resultant wave is not sinusoidal; various types of resultant waveform are discussed in this and the subsequent sections. Before detailed consideration of these syntheses, however, a general property of sine-waves in combination will be demonstrated.

Fig. 9*a* shows a fragmentary portion of two sine-waves A and B, whose equations may without loss of essential generality be expressed as

$$y_1 = a \sin x,$$
$$y_2 = b \sin (px + \phi),$$

respectively, so that the wavelength of A is p times the wavelength of B, and when $x = 0$ the value of y_2 is $b \sin \phi$. Peaks (i.e. crests and troughs) in the resulting waveform $y = y_1 + y_2$ occur where $dy/dx = 0$, i.e.

$$a \cos x + bp \cos (px + \phi) = 0;$$

or

$$a \cos x = -bp \cos (px + \phi).$$

In Fig. 9*b* the curves A′, B′ represent dA/dx, $-dB/dx$, respectively, i.e. the waves $a \cos x$ and $-bp \cos (px + \phi)$. It is evident that in the range PQ there will be only one peak in the resultant waveform, as shown in Fig. 9*c* (the peak in this case being a crest), as there is only one intersection in this range in Fig. 9*b* to fulfil the condition $dy/dx = 0$. The "intuitive" idea that there should be two crests in the corresponding range in the resulting waveform, one caused by the crest C_1 in A and the other by the crest C_2 in B, Fig. 9*a*, is thus seen to be erroneous.

Detailed investigation, too involved to be included in this present work, shows that there can never be more peaks in a resultant waveform than there are in the component wave of highest frequency present; in many cases there are less.

Beating. When two sine-waves, the ratio of the frequencies of which is nearly unity, are added together, the phenomenon of beating occurs; the resultant wave is not sinusoidal, but has the appearance of a sine-wave with an approximately sinusoidal variation in amplitude.

Let the two waves be

$$y_1 = a \,.\, \cos \omega t,$$
$$y_2 = b \,.\, \cos (\omega + \Delta\omega)t,$$

where $\Delta\omega$ may be positive or negative but $(1 + \Delta\omega/\omega)$ is nearly equal to unity, i.e. $\Delta\omega/\omega$ is small. If, as before, $y = y_1 + y_2$, then

$$\begin{aligned} y &= a \,.\, \cos \omega t + b \,.\, \cos (\omega + \Delta\omega)t \\ &= (a + b \,.\, \cos \Delta\omega t) \cos \omega t - b \,.\, \sin \Delta\omega t \,.\, \sin \omega t \\ &= r \,.\, \sin (\omega t + \psi), \end{aligned}$$

$$\left.\begin{aligned} &\text{where} \quad r^2 = (a + b \,.\, \cos \Delta\omega t)^2 + (b \,.\, \sin \Delta\omega t)^2 \\ &\text{i.e.} \quad r = \sqrt{a^2 + b^2 + 2ab \,.\, \cos \Delta\omega t}. \end{aligned}\right\} \quad . \quad . \quad (6.1)$$

The amplitude r of the resultant wave thus varies between the limits $(a + b)$ and $(a - b)$ with the frequency $\Delta\omega/2\pi$ C.P.S. (since $\Delta\omega$ is an angular velocity in radians per second and there are 2π radians in one revolution of a generating vector—i.e. in one cycle) and this frequency is the difference between the frequencies of the component waves :

$$\Delta\omega/2\pi = (\omega + \Delta\omega)/2\pi - \omega/2\pi,$$

and is known as the *beat frequency*.

Fig. 10 illustrates the addition of the two waves $2 \,.\, \cos 5t$ and $\cos 6t$. In this case $\Delta\omega/2\pi = 1/2\pi$, and the beat period is 2π radians or 360°. In this period there are 5 crests and 5 troughs

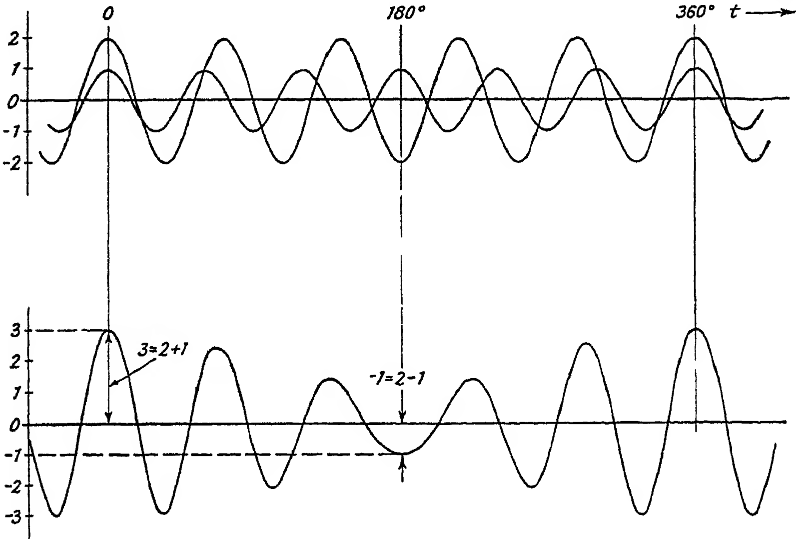


FIG. 10.—Beats : the wave $2 \cos 5t + \cos 6t$.

of the resultant wave, and the amplitude at the beat maxima is $2 + 1 = 3$, while the amplitude at the beat minima is $2 - 1 = 1$.

The resultant wave cannot strictly be regarded as a sine-wave with a sinusoidal variation in amplitude, since the expression for the amplitude in (6.1) involves the extraction of the square root of the quantity $a^2 + b^2 + 2ab \,.\, \cos \Delta\omega t$, and unless $a = b$ the result is not sinusoidal. Moreover, there is the phase-angle ψ to be taken into account. This angle is not constant but varies throughout the beat cycle (i.e. throughout the range from one beat maximum to the next). In fact,

$$\tan \psi = -\frac{a + b \,.\, \cos \Delta\omega t}{b \,.\, \sin \Delta\omega t} \quad . \quad . \quad . \quad . \quad (6.2)$$

This variation in the phase-angle does not affect the result that the beat frequency is the difference between the component frequencies, but may affect the apparent frequency of the resultant distorted wave—i.e. the number of peaks of the wave in a beat cycle. Thus in the case considered above, there should be $|\omega/\Delta\omega|$ crests per beat cycle (the sign $|\ |$ indicating that the numerical value is to be taken without sign) and this is true if $a > b$; but if $a < b$ the effect of the varying phase-angle is to provide another crest per beat if $\Delta\omega$ is positive and to remove a crest per beat if $\Delta\omega$ is negative. The practical result may be stated thus:

The resultant wave has the same apparent frequency as the major * component (i.e. that with the greater amplitude) and its amplitude varies between the sum and the difference of the component amplitudes, the beat frequency being the difference between the frequencies of the components.

In Fig. 11 the addition of the waves $2 . \cos 5t$ and $3 . \cos 6t$ is illustrated, and there are now 6 crests in the beat cycle, since

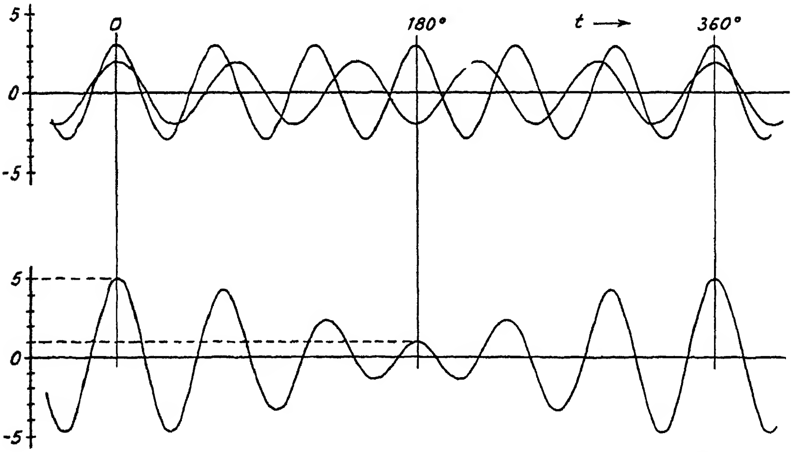


FIG. 11.—Beats: The wave $2 \cos 5t + 3 \cos 6t$; compare Figs. 10 and 14.

$3 . \cos 6t$ is the major component. The maximum amplitude is $3 + 2 = 5$, and the minimum amplitude $3 - 2 = 1$.

* The terms " major " and " minor " are widely used with this and the opposite connotation. It should be noted that the so-called " major " and " minor " harmonics of the torque output of an internal combustion engine may or may not be respectively the major and minor components of a recorded waveform, according to resonance conditions and the engine design details.

Note.—Actually the criterion which determines the number of peaks in a beat cycle is more complicated than that quoted, although this is sufficiently accurate in most practical cases. Confining attention to the case where $\Delta\omega$ is positive and $\omega/\Delta\omega$ is integral, the two curves can be put in the form $a \,.\, \cos kx$ and $b \,.\, \cos (k + d)x$, where k and d are integral and positive, by a suitable choice of the new variable x in terms of t. In the range of x from 0 to 2π radians there are k cycles of $a \,.\, \cos kx$, $(k + d)$ of $b \,.\, \cos (k + d)x$, and d beat cycles. In one beat cycle there are thus k/d cycles of the lower frequency wave $a \,.\, \cos kx$, and $(k + d)/d = k/d + 1$ of the higher frequency wave $b \,.\, \cos (k + d)x$. Peaks in the resultant wave occur at maxima and minima of the function $y = a \,.\, \cos kx + b \,.\, \cos (k + d)x$; these stationary values occur where $dy/dx = 0$, i.e. where

$$ak \,.\, \sin kx + b(k + d) \sin (k + d)x = 0.$$

Consider now the graphs of $ak \,.\, \sin kx$ and $-b(k + d) \sin (k + d)x$ over half a beat cycle. These are indicated in Fig. 12 for the particular case

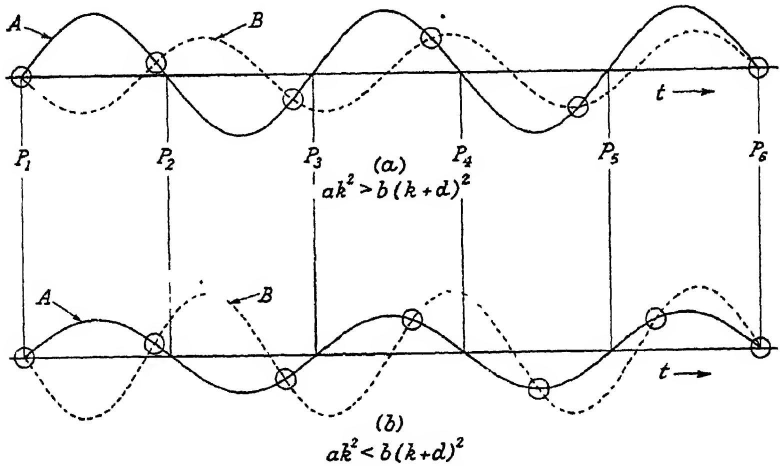


FIG. 12.—Number of peaks in a beat cycle.

$k/d = 5$. The zero points of the function dy/dx are given by the intersections of the two curves $ak \,.\, \sin kx$ and $-b(k + d) \sin (k + d)x$, which for convenience will be referred to as A and B, respectively. Two cases are illustrated in Fig. 12. First (Fig. 12*a*), if $b(k + d)$ is sufficiently small compared with ak, there are intersections at P_1 and P_6, and one intersection in each of the $(k/d - 1)$ ranges P_1P_2, P_2P_3, P_3P_4, and P_4P_5, where P_1, P_2 . . . P_6 are successive zero-points of the curve A as shown in the diagram. There is no intersection in the last range P_5P_6. Thus in all there are $2k/d$ intersections in the beat cycle (only one of the end-points being counted) and half of these represent maxima in the waveform, the other half minima. There are thus k/d crests and k/d troughs per beat cycle. If, however, $b(k + d)$ is sufficiently large compared with ak (Fig. 12*b*), the dotted curve B lies partly outside the curve A in the range P_5P_6, and there is an extra intersection per half-cycle of the beat. In all there are now $2(k/d + 1)$ intersections in the beat cycle, representing $(k/d + 1)$ crests and $(k/d + 1)$ troughs in the waveform.

The criterion is the comparative slope of the two curves in Fig. 12 at the point P_6. If the slope of B is numerically greater than that of A at P_6, then Fig. 12*b* is the appropriate diagram; and, if less, Fig. 12*a*. Numerically the gradients of A and B are ak^2 and $b(k+d)^2$, respectively, and hence :

If $ak^2 < b(k+d)^2$, there are $(k/d+1)$ crests per beat cycle, and
if $ak^2 > b(k+d)^2$, there are k/d crests per beat cycle.

It follows that if $b > a$, i.e. the higher frequency component is major, the resultant wave has the same apparent frequency as the major component ; this is the case in the example illustrated by Fig. 11. If, however, $a > b$, so that the lower frequency wave is the major, the apparent frequency of the resultant wave is that of the minor component unless $ak^2 > b(k+d)^2$. In Fig. 10 this last condition is fulfilled, and the frequency of the major component " shows up " in the resultant waveform.

If $ak^2 = b(k+d)^2$, three stationary values coincide and there is a " spread trough " at the waist of the beat (i.e. where the amplitude r is minimum).

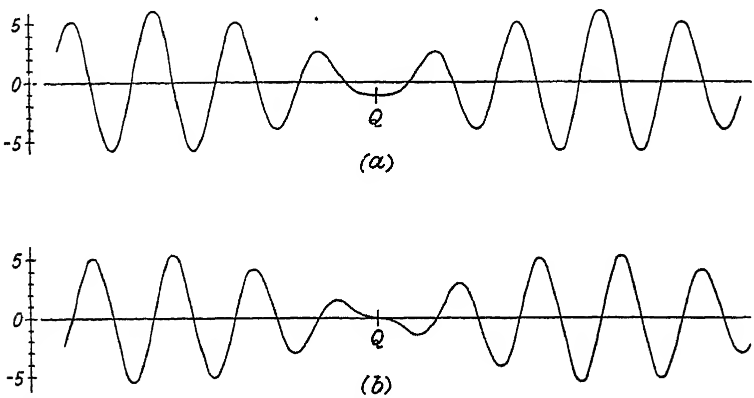


FIG. 13.—Critical cases of beats : (*a*) " spread trough " at waist ; (*b*) inflection at waist.

This case is illustrated in Fig. 13*a*, which shows the graph of the function $3.6 \cos 5t + 2.5 \cos 6t$; the spread trough is marked at Q.

This criterion holds for the addition of two cosine waves—i.e., two sine-waves which are so related that at one point in each cycle of the resultant wave the two components " peak " together, and the resulting wave is symmetrical. In the case of two sine-waves which have zero values together —as, for example, $a \,.\, \sin kx$ and $b \,.\, \sin (k+d)x$ which are both zero at $x = 0$ and $x = \pi$—another criterion is applicable : if $ak = b(k+d)$ the two components have gradients, at the waist of the beat, which are equal in magnitude but opposite in sign, so that there is an inflection in the resultant waveform at this point. The conditions thus are

If $ak > b(k+d)$, there are k/d crests per beat cycle, and
if $ak < b(k+d)$, there are $(k/d+1)$ crests per beat cycle.

Fig. 13*b* illustrates the critical condition in the particular case where $k = 5$ and $d = 1$: the graph shown is that of the function $3 \sin 5t + 2.5 \sin 6t$ and the inflection is marked at Q.

7. Separation of peaks in beating waveforms.

Fig. 14 shows the waveform $2 \cos 6t + \cos 7t$, and a comparison of this diagram with Fig. 11, which represents the function $2 \cos 5t + 3 \cos 6t$, shows that in both cases there are 6 crests in the beat cycle. Given a waveform containing 6 crests per beat cycle, there is no evidence, according to the properties so far considered, to indicate whether the minor component is of higher or

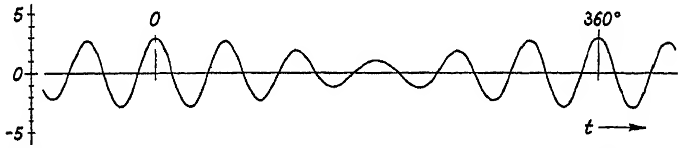


FIG. 14.—Beats: the wave $2 \cos 6t + \cos 7t$; compare Fig. 11.

lower frequency. There is, however, one property of beating waveforms which affords the information required. Referring to the part of the beat cycle where the amplitude is a maximum by the convenient name "bulge," and to the part where the amplitude is a minimum by the name "waist," the rules are as follows:

(i) If the minor component is of higher frequency than the major, the distance between successive peaks at the bulge is less than the corresponding distance at the waist.

(ii) If the minor component is of lower frequency than the major, the distance between successive peaks at the bulge is greater than the corresponding distance at the waist.

The significant parts of the diagrams, Figs. 11 and 14, are replotted in Fig. 15 to illustrate these rules. S_B and S_W are the peak separations (trough to trough) at bulge and waist respectively.

Note.—The mathematical basis for the above rules is easily demonstrated. Let the two waves be $a \,.\, \cos kx$ and $b \,.\, \cos (k + d)x$ as before, but now let it be postulated that $a > b$, while d may be positive or negative. Thus $a \,.\, \cos kx$ is the major component, and if d is positive rule (i) is to be proved, while if d is negative (ii) is the appropriate rule. The value $x = 0$ corresponds to the bulge of the beat, while $x = \pi$ corresponds to the waist. The first trough of the major component occurs at $x = \pi/k$, and for this value the slope of the function $y = a \,.\, \cos kx + b \,.\, \cos (k + d)x$ is given by

$$\begin{aligned} dy/dx &= -[ak \,.\, \sin \pi + b(k + d) \sin (\pi + d\pi/k)] \\ &= b(k + d) \sin (d\pi/k). \end{aligned}$$

If d is positive, the slope is positive also, and thus at $x = \pi/k$ the curve is rising; thus the trough has been passed, and $S_B < 2\pi/k$. If, on the other

hand, d is negative, similar reasoning shows that $S_B > 2\pi/k$. At the waist, the waveform can be represented by the function

$$\pm y^* = a \,.\, \cos kx - b \,.\, \cos (k + d)x,$$

since here the minor component has a trough where the major has a crest, or vice versa; it is easy to show that if d is positive $S_W > 2\pi/k$, while if d is negative $S_W < 2\pi/k$. The rules (i) and (ii) are thus established.

8. Beat envelopes.

If curves are drawn as smoothly as possible on either side of a beating waveform, one touching at or near each of the crests and the other at or near each of the troughs, as shown by the dotted lines in Fig. 16*a*, these curves are termed *envelopes* and together form what is known as the "primary envelope characteristic," or more simply as the "first envelope." When the beating waveform

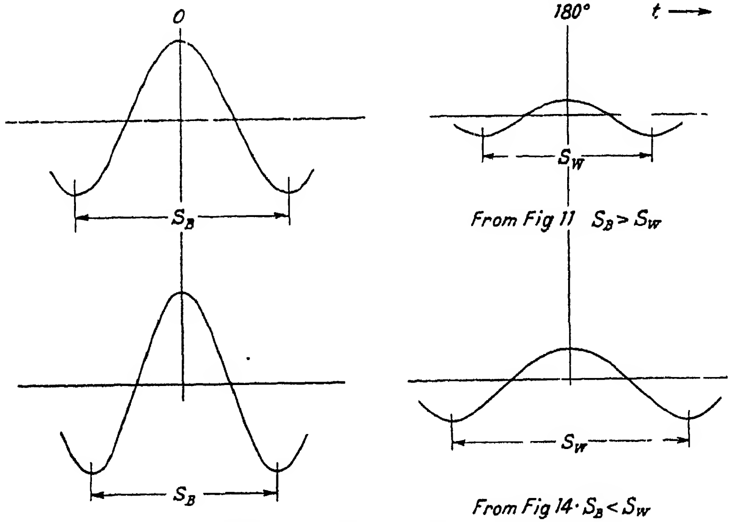


FIG. 15.—Peak separation in beating waveforms.

has but two components, the bottom envelope is the "mirror-image," in the centre-line, of the top envelope. This fact is illustrated in Fig. 17, in which the curve of Fig. 14 is replotted as though it were "rectified" in the electrical sense—i.e. those parts P′ of the curve which ordinarily fall below the centre-line are plotted with a change of sign. It is seen that the same envelope touches these rectified peaks as touches the true crests P.

Although the envelopes are not truly sinusoidal, it is convenient to imagine them to be so for the purpose of the statement that the top and bottom envelopes are anti-phased. Also, referring to the

* The positive sign is taken if the major component has a crest at the waist, and the negative sign if it has a trough thereat.

area between the envelopes as the envelope "strip" (shown shaded in Fig. 16*b*) the strip width at the bulge is equal to the sum of the double amplitudes of the components, while the strip width at the waist is equal to the difference between the double amplitudes of the components. Thus the total variation in height of either

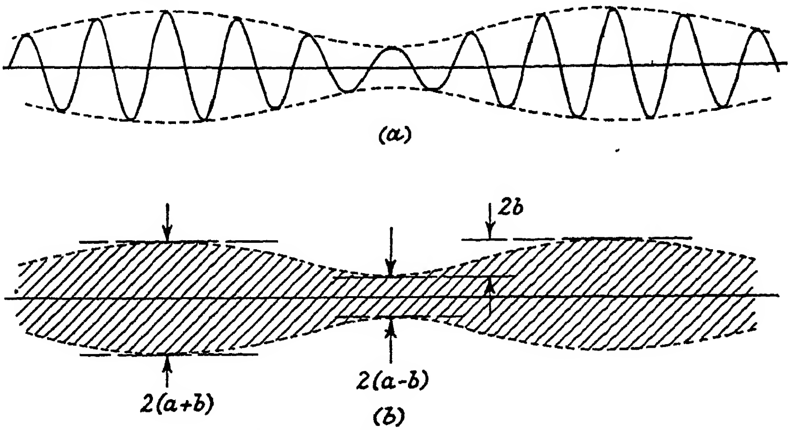


FIG. 16.—Beat envelopes.

envelope is equal to the double amplitude of the minor component, and it is convenient to use the term "amplitude" with reference to the envelope as though it were sinusoidal; it is then said that the amplitude of the envelope is the amplitude of the minor component.

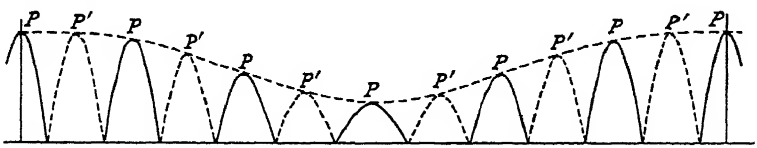


FIG. 17.—Symmetry of beat envelopes: broken-line portions, which have suffered a reversal of sign, touch the same envelope as the full-line portions.

Note.—At first sight it may appear that this last property depends upon the fact that the component waves in Fig. 14 are integrally related cosine waves—i.e. their crests coincide at the bulge, and the crest of the major coincides with the trough of the minor at the waist. Actually, however, it does not matter what the phase-relation may be; so long as the beat cycle contains a reasonably large number of cycles of the major component—say, more than three—if the condition is observed that the envelopes shall be drawn "as smoothly as possible" the rule holds good. It would require more space than is available to demonstrate this fact adequately, but experience shows it to be the case. It is suggested that the reader satisfies himself upon this point by drawing the waveforms which result from the addition of two waves with a suitable frequency ratio, and with various

phase-relationships. The case of the addition of two waves whose frequency ratio is of the form 1 : 2 or 2 : 3 requires special consideration, as these waveforms are examples of a class which is intermediate between beating and the type of waveform which will next be considered.

9. Summary of beat properties.

The properties of beating waveforms described in the foregoing sections 6-8 are recapitulated below for convenience.

When two waves, whose frequency ratio is nearly unity, are added together, the resultant waveform possesses the following characteristics :

(i) There is a periodic fluctuation in the amplitude of the wave, which is said to be beating.
(ii) The top and bottom envelopes are anti-phased.
(iii) The frequency of the envelope is equal to the difference between the frequencies of the components.
(iv) The greatest width of the envelope strip is equal to the sum of the double amplitudes of the components.
(v) The amplitude of the envelope is equal to the amplitude of the minor component.
(vi) In a given length of the resultant wave the number of peaks is, except in certain critical cases detailed in Section 6, the same as in the same length of the major component wave.

10. Addition of two waves of high frequency-ratio.

The next type of waveform to be discussed is that resulting from the addition of two waves, the ratio of whose frequencies differs greatly from unity. Postponing treatment of the cases where the ratio is of the order of 1 : 2 or 1 : 3, consider the waveform shown in Fig. 18a, which represents the function

$$y = a_1 \sin x + a_6 \cos 6x,$$

where a_1 and a_6 have the particular values 2, 1 respectively. The frequencies of both the components are apparent in the waveform ; there are 6 crests (and 6 troughs) in the cycle, corresponding to the high-frequency component $a_6 \cos 6x$, and the wave has a low-frequency surge which " follows " the low-frequency component $a_1 \sin x$ shown by the dotted line.*

* The terms " low-frequency " and " high-frequency " are used, in waveform analysis, in a relative sense ; more strictly they should be replaced by " lower- " and " higher- frequency," but in any case the distinction between low and high frequencies is one of relativity : a frequency of 100,000 C.P.M. would be accounted " high " by a mechanical engineer, and " low " by an electronic engineer.

In Fig. 18*b* the curve is repeated with the top and bottom envelopes drawn in. These envelopes are both parallel to the low-frequency component wave, and are thus in phase. Moreover, at the crest P_1 of the low-frequency component the height P_1Q_1 is given by

$$P_1Q_1 = a_1 \sin 90° + a_6 \cos (6 \times 90°)$$
$$= a_1 - a_6,$$

while the height P_2Q_2 at the trough of the low-frequency component is given by

$$P_2Q_2 = a_1 \sin 270° + a_6 \cos (6 \times 270°)$$
$$= -(a_1 + a_6)$$

(being negative since P_2 is below the centre-line).

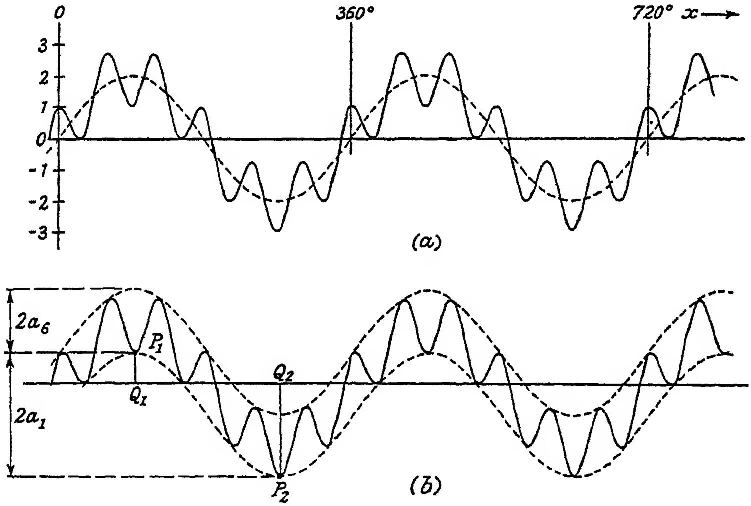


FIG. 18.—Addition of two waves of high frequency-ratio: the wave $a_1 \sin x + a_6 \cos 6x$.

Thus the envelope amplitude equals the amplitude of the low-frequency component as

$$(a_1 - a_6) - [-(a_1 + a_6)] = 2a_1.$$

Either envelope therefore represents the low-frequency component in both amplitude and phase; and measurement shows that the envelope strip-width (i.e. the constant vertical distance between the two envelopes) equals the double amplitude of the high-frequency component.

Figs. 19, 20 show two further examples of this type of waveform. In Fig. 19 the graph represents the function $5 \cos 4x + 5 \cos 15x$, the low-frequency component being drawn dotted on the same

centre-line. There are 15 crests and 15 troughs in a cycle, corresponding to the high-frequency component, and the curve follows the general trend of the low-frequency component. In Fig. 20 the function represented is $5 \sin x + 8 \cos 12x$.

The general characteristics of this type of waveform may be recapitulated thus :

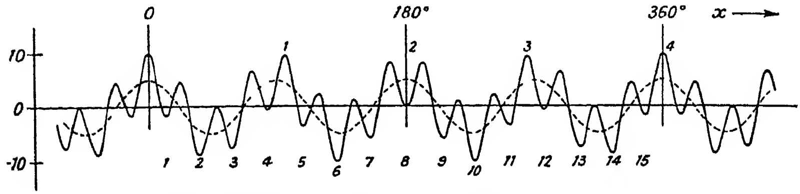


Fig. 19.—High frequency-ratio : the wave $5 \cos 4x + 5 \cos 15x$.

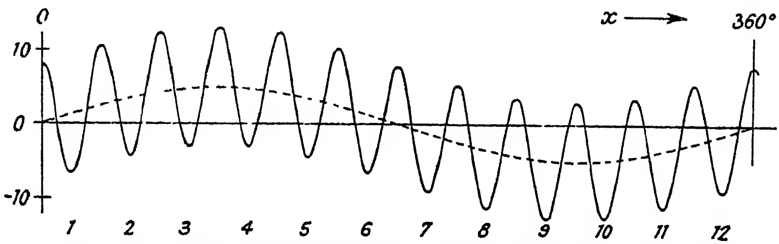


Fig. 20.—High frequency-ratio : the wave $5 \sin x + 8 \cos 12x$.

(i) The high-frequency component appears in the resultant wave as a "ripple" superimposed on the low-frequency component.

(ii) Both envelopes represent, in amplitude and phase, the low-frequency component.

(iii) The width of the envelope strip is equal to the double amplitude of the high-frequency component.

The reader is recommended to experiment with examples of this synthesis to convince himself on the generality of these properties. (See Section 13.)

Occurrence. In records of mechanical vibration taken by means of vibrographs, it frequently happens that the natural vibration of the instrument is excited by some agency, thus adding a corresponding component into the recorded waveform. If this phenomenon is anticipated, the instrument can be so designed that there is a high-frequency ratio between the physical variation being recorded and the natural vibration of the vibrograph, so that in the case where the physical variation is sinusoidal the resulting

waveform is of the type discussed in this section; as will be seen in Chapter IV, the two components can easily be analysed out, thus enabling the true physical variation to be obtained from the record. If possible, the natural frequency of the instrument is put well below the range of frequencies which are expected to be recorded, so that if the excitation is non-sinusoidal resonance will not be encountered with higher harmonics. Ker Wilson shows an example of a recorded waveform wherein a definite component at the natural frequency of the instrument is clearly visible (reference 4 in the Bibliography at the end of the book).

The same type of effect is met with in electrical work, or in the use of electrical apparatus recording mechanical vibrations, where a "mains hum" component at 50 C.P.S. or 3000 C.P.M. makes its appearance on the record owing to faulty design of the equipment. Dependent on the frequency of the vibration being recorded,

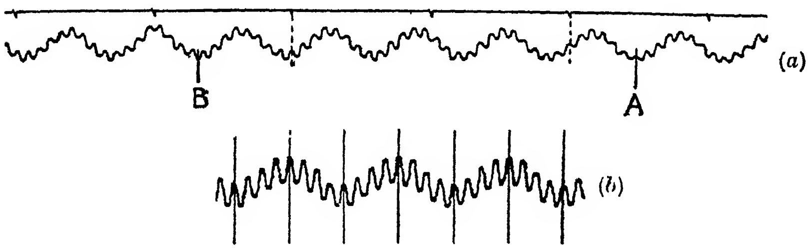


FIG. 21.—Mains hum: (*a*) timing marks at 1/4 second intervals, mains hum is high-frequency ripple; (*b*) vertical lines at 1/100 second intervals, mains hum is low-frequency surge.

the hum may appear either as a low-frequency surge or as a high-frequency ripple. Fig. 21 shows examples of both types. At (*a*), the upper trace indicates time-intervals of 1/4 second, and it can be seen that in the interval of 1/2 second indicated by the two broken vertical lines there are 25 cycles of the high-frequency ripple, which is therefore at mains frequency. At (*b*), the vertical lines indicate time-intervals of 1/100 second, and it is seen that the low-frequency surge in this trace is at mains frequency.

11. Intermediate frequency ratio.

The types of synthesis so far discussed have concerned component waves of two classes: (*a*) with a frequency ratio nearly equal to unity, and (*b*) with a frequency ratio greatly different from unity. Two special cases of an intermediate type are waveforms in which the frequency ratio of the two component waves is exactly two or three; some examples of these forms are given in Figs. 22,

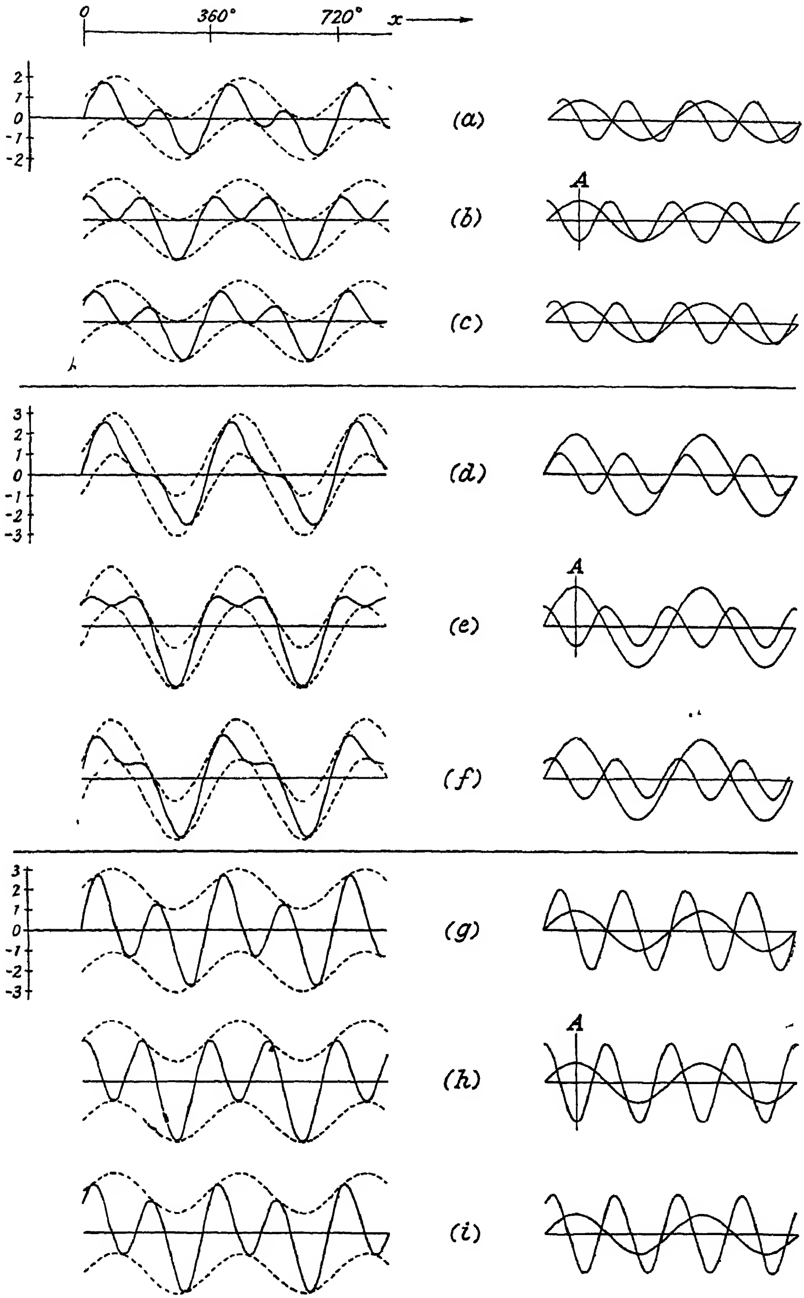


FIG. 22.—Addition of two waves with frequency-ratio 2 : 1.

23. Apart from these special cases, most other examples of wave-forms obtained by the addition of two waves, and not included in either of the classes (*a*) or (*b*), have characteristics which are typified by the diagram Fig. 24. These various waveforms are now discussed in turn.

Frequency ratio = 2. Nine special cases of the general form

$$y = a_1 \sin x + a_2 \sin (2x + \phi) \qquad . \qquad . \quad (11.1)$$

are depicted in Fig. 22. The component waves are shown on the right-hand side of the diagram; in Figs. 22*a*, *b*, *c* the two components are of equal amplitudes, $a_1 = a_2$; in Figs. 22*d*, *e*, *f* the amplitude of the low-frequency component is double that of the high-frequency component, $a_1 = 2a_2$, and in the last three examples the amplitude of the high-frequency component is double that of the low-frequency component, $a_2 = 2a_1$. In the first example in each trio, the phase-angle ϕ is zero, so that Figs. 22*a*, *d*, *g* represent respectively the functions

$$\sin x + \sin 2x,$$
$$2 \sin x + \sin 2x,$$
$$\sin x + 2 \sin 2x.$$

The resulting waves are skew-symmetrical, since both the components are skew-symmetrical about the values $x = 0, \pi, 2\pi$, etc.

In the second example in each trio, Figs. 22*b*, *e*, *h*, the phase-angle is 90°, so that the graphs represent respectively the functions

$$\sin x + \cos 2x,$$
$$2 \sin x + \cos 2x,$$
$$\sin x + 2 \cos 2x.$$

If the basic variable x is changed to x_1, where

$$x_1 = x - 90°,$$

then, since $\quad \sin x = \sin (x_1 + 90°) = \cos x_1,$

and $\quad \cos 2x = \cos (2x_1 + 180°) = - \cos 2x_1,$

it is evident that the resulting waveform is symmetrical, as both $\cos x_1$ and $\cos 2x_1$ are symmetrical, about the values $x_1 = 0$ (indicated by the letter A in the right-hand part of the diagram), π, 2π, etc.

In the remaining examples the phase-angle ϕ is given the value 45°, and the waveforms are neither symmetrical nor skew-symmetrical.

It is important to note that in every case the overall width of the resultant wave is appreciably less than the sum of the double amplitudes of the components.

This discrepancy is greatest in the second wave in each trio, and the greater the ratio a_1/a_2 becomes the smaller is the ratio (total

width of resultant wave/sum of double amplitudes of components). This fact is of the utmost importance in engineering vibration analysis, when structural components are subjected to vibrational loads at two such frequencies simultaneously. Suppose, for example, that a periodic stress of amplitude 8000 lb./in.^{2} is produced in a certain part of a machine at a frequency of 2600 C.P.M., and simultaneously a periodic stress of amplitude 4000 lb./in.^{2} is produced at the same place at a frequency of 5200 C.P.M. If these two sinusoidal stress components have a phase-relationship similar to that depicted in Fig. 22*e*, the resultant total stress does not vary between the limits $\pm$ 12,000 lb./in.^{2}, but between the limits $+$ 6000 and $-$ 12,000 lb./in.^{2}, the waveform produced by a recording strain gauge being similar to Fig. 22*e*. Such a phenomenon must necessarily have an effect upon the fatigue life of the machine.

The broken lines on the left-hand part of Fig. 22 represent the low-frequency component, displaced vertically so that the vertical distance between the two broken lines on each waveform is equal to the double amplitude of the high-frequency component, and the two lines are symmetrically disposed with reference to the centre-line. It may be observed that these lines constitute envelopes, between which the waveforms lie, touching them at or near each crest and trough. Both the top and bottom envelopes assume the only form possible if they are to be sinusoidal, of equal amplitudes, and mutually in-phase, although if these conditions are neglected it is possible to construct enveloping lines which may have amplitudes very much smaller than the amplitude of the low-frequency component; this is particularly true of the waveforms shown in Figs. 22*b*, *e*, *h*, in each of which all the crests lie on a straight line parallel to the centre-line. This complication might appear to render inapplicable a scheme of *analysis*, based upon envelope properties, to this species of waveform; it is shown in Chapter IV (p. 99), however, that in fact a considerable measure of arbitrary choice of envelopes is permissible without invalidating the method.

It is to be noted that in Fig. 22*d* there are inflections on the centre-line. The presence of the high-frequency component is still indicated by the distortion of the wave from the pure sine-wave shape, but the small crests and troughs seen in Fig. 22*a* have now coalesced into points of inflection. It is easy to show that if the amplitude of the low-frequency component is greater than twice the amplitude of the high-frequency component (i.e. when $a_1 > 2a_2$), and the waves have this particular phase-relationship $\phi = 0$, there is not even a point of inflection, where the tangent is parallel to the centre-line, but merely a point of oblique contra-

flection; such a condition is represented approximately in the part BC of the curve in Fig. 24a.

The mathematical proof of this fact is simple. Since

$$y = a_1 \sin x + a_2 \sin 2x,$$

the points where the tangent is parallel to the centre-line are given by

$$dy/dx = a_1 \cos x + 2a_2 \cos 2x = 0,$$

i.e. $$a_1 \cos x + 2a_2(2 \cos^2 x - 1) = 0,$$

whence $$4a_2 \cos^2 x + a_1 \cos x - 2a_2 = 0,$$

i.e. $$\cos x = -\frac{a_1 \pm \sqrt{a_1^2 + 32a_2^2}}{8a_2},$$

Since $\sqrt{a_1^2 + 32a_2^2} > a_1$, there is always determined at least one value of $\cos x$, giving two values of x; but if $a_1 > 2a_2$,

$$a_1 + \sqrt{a_1^2 + 32a_2^2} > 2a_2 + \sqrt{(2a_2)^2 + 32a_2^2},$$

i.e. $> 8a_2$, so that $\cos x < -1$. Thus if $a_1 > 2a_2$ only two real values of x are given, and there is only one crest and one trough per cycle.

The same type of phenomenon occurs in the case where $\phi = 90°$, i.e. where the wave is of the form

$$a_1 \sin x + a_2 \cos 2x.$$

With the change of variable $x = x_1 + 90°$ already utilised in connection with this case, the wave can be represented as

$$y = a_1 \cos x_1 - a_2 \cos 2x_1,$$

and the condition for a tangent parallel to the centre-line is

$$a_1 \sin x_1 - 2a_2 \sin 2x_1 = 0,$$

i.e. $$\sin x_1 (a_1 - 4a_2 \cos x_1) = 0.$$

Whatever values are assigned to a_1 and a_2, there are turning points at $x_1 = 0$ and $180°$, given by $\sin x_1 = 0$, and if $a_1 \leqslant 4a_2$ there are turning points at two other values of x_1 given by $\cos x_1 = a_1/4a_2$; but if $a_1 > 4a_2$ no real values of x_1 are determined by this condition. In this case there is a "spread peak" similar to that at Q in Fig. 13a.

For other values of the phase-angle ϕ there are different criteria; these are easily determined numerically in particular cases.

Frequency ratio = 3. In Fig. 23 are depicted three particular cases of the general form

$$y = a_1 \sin x + a_3 \sin (3x + \phi) \quad . \quad . \quad (11.2)$$

with $a_1 = a_3 = 1$.

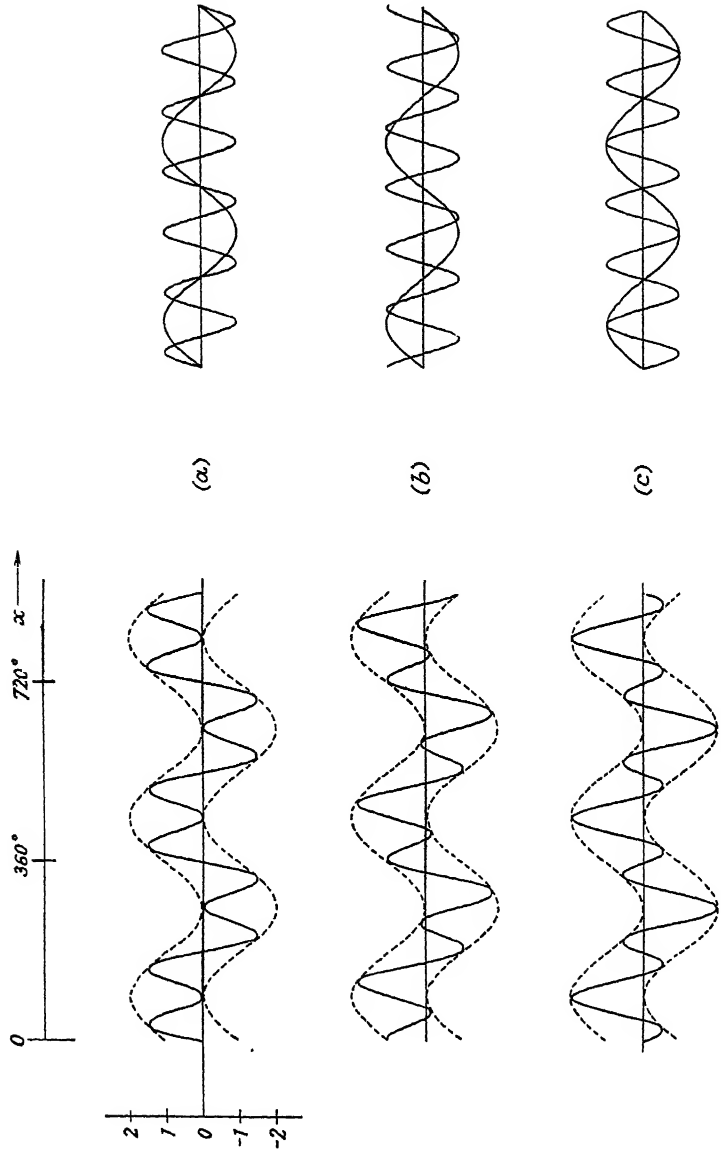


Fig. 23.—Addition of two waves with frequency-ratio 3 : 1.

In Figs. 23a, b, c the phase-angle ϕ is given the values 0, 90°, 180° respectively, so that the individual waveforms are

$$\left.\begin{aligned}&(a)\ y = \sin x + \sin 3x.\\&(b)\ y = \sin x + \sin (3x + 90°) = \sin x + \cos 3x.\\&(c)\ y = \sin x + \sin (3x + 180°) = \sin x - \sin 3x.\end{aligned}\right\} \quad . \qquad (11.3)$$

From a study of the right-hand part of the diagram, where the component waves are drawn, it may be seen that except in case (c), i.e. where $\phi = 180°$, the total width of the waveform must be less than the sum of the double amplitudes of the components, for it is only in this case that both components reach maximum values, and both reach minimum values, simultaneously. In Fig. 23*a*, the total width is approximately three-quarters of the sum of the component double amplitudes.

All the examples of the form (11.2) illustrated in Fig. 23 have component waves of equal amplitudes, $a_1 = a_3 = 1$. The effect

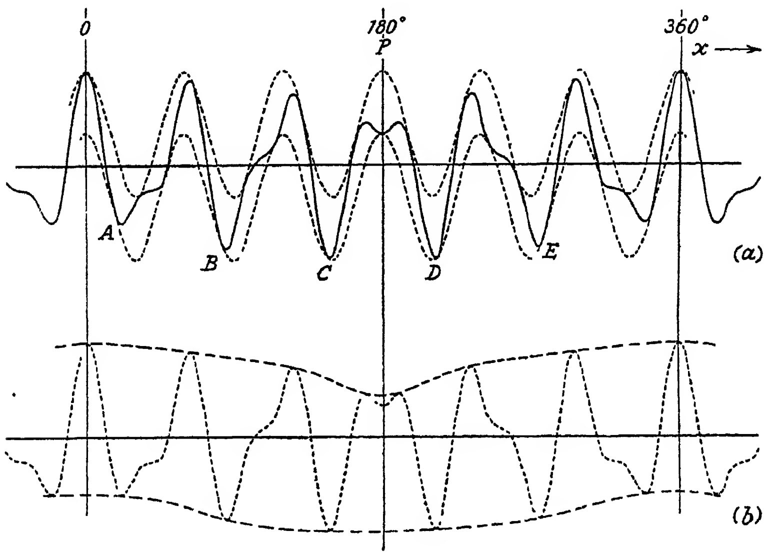


FIG. 24.—Apparent low-frequency surge in special type of waveform where one frequency is nearly a multiple of the other.

of varying the amplitude ratio a_1/a_3 is similar to the effect of varying the ratio a_1/a_2 in Fig. 22, and need not here be enlarged upon.

The waveforms depicted in Fig. 23 will doubtless be familiar to students of advanced electrical design, particularly in connection with transformer characteristics. The output from a transformer fed with a 50 C.P.S. pure sine-wave input does frequently contain quite a large component at $150 = 3 \times 50$ C.P.S. frequency.

Frequency ratio 11 : 6. Fig. 24 illustrates an example of the last intermediate type of two-component waveform which will here be discussed. The wave represents the function

$$y = 2 \cos 6x + \cos 11x. \qquad (11.4)$$

The broken lines show the low-frequency component drawn in the same manner as in Fig. 22. The frequency ratio being nearly equal to two, various parts of the waveform approximate in some degree to Figs. 22*d*, *e*, *f*. Thus the part CD is of the same general shape as a cycle of Fig. 22*e*, and DE corresponds approximately to a cycle of Fig. 22*d*; similarly, the mirror-image, in the centre-line, of the part AB of Fig. 24*a* shows the same general form as a cycle of Fig. 22*f*.

The mathematical reason for this property is easy to substantiate. As

$$\begin{aligned} \cos 11x &= \cos (12x - x) \\ &= \cos (12x - \phi), \text{ where } \phi = x, \\ y &= 2 \cos 6x + \cos (12x - \phi). \end{aligned}$$

Changing the variable x to x_1, where $x_1 = 6x$,

$$\left.\begin{aligned} y &= 2 \cos x_1 + \cos (2x_1 - \phi) \\ \text{where} \quad \phi &= x_1/6. \end{aligned}\right\} \quad . \quad (11.5)$$

In the part CD of the wave, corresponding to the range $150° \leqslant x \leqslant 210°$, the range of x_1 is $900° \leqslant x_1 \leqslant 1260°$, or, subtracting $1080° = 3 \times 360°$,

$$-180° \leqslant x_1 \leqslant 180°.$$

Also, ϕ at C is given by

$$\phi_C = x_C = 150° = 180° - 30°,$$

and at D $\qquad \phi_D = x_D = 210° = 180° + 30°.$

At C, (11.5) can therefore be written

$$y_C = 2 \cos x_1 - \cos (2x_1 + 30°),$$

and at D, $\qquad y_D = 2 \cos x_1 - \cos (2x_1 - 30°).$

Also, at P, midway between C and D ($x = 180°$),

$$y_P = 2 \cos x_1 - \cos 2x_1.$$

From C to P, and from P to D, the phase-angle ϕ has a variation of only 30°; and it has already been shown that by a suitable change of variable the function represented by Fig. 22*e* can be put in the form $y = 2 \cos x_1 - \cos 2x_1$. Thus such discrepancies as may exist between the shape of the part CD and the shape of a cycle of Fig. 22*e* are those small differences due to the slight variation in the phase-angle ϕ over the range 30° to − 30°. Similar reasoning is applicable to the other parts of the waveform.

Fig. 24*b* shows, in broken line, the effect of drawing in a pair of "envelopes" touching the waveform at or near the crests or troughs of the low-frequency component; there is an apparent surge whose cycle extends over the range from $x = 0$ to $x = 2\pi$, i.e. the frequency of the surge is the same as that of $\cos x$ in this

example. In general, when the frequency F_1 of one component is nearly equal to a multiple KF_2 of the frequency F_2 of the other component, there is an apparent surge whose frequency is given by the difference $F_1 \sim KF_2$.* This phenomenon is termed the "apparent low-frequency surge effect."

12. Three components.

The number of different types of waveform arising from the addition of three component sinusoidal waves is so large as to preclude any possibility of discussing them all in detail. In this section two such types only are illustrated, to draw attention to certain general principles. Commencing with a two-component

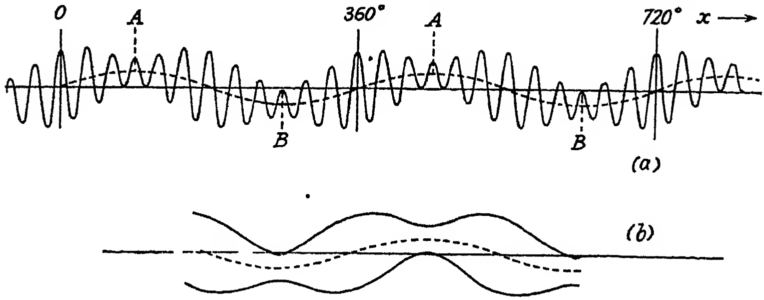


FIG. 25.—Three-component wave, with low-frequency surge.

waveform, and adding thereto a third component, the two types here considered are

(i) third component of much lower frequency than the other two components, and
(ii) third component of much higher frequency than the other two components.

(i) *Third component of much lower frequency.*

Fig. 25*a* illustrates the function

$$y = 5 \sin x + 8 \cos 12x + 4 \cos 10x. \qquad (12.1)$$

The third component, $5 \sin x$, is shown in broken line on the same centre-line. If the graph of the function

$$y = 8 \cos 12x + 4 \cos 10x, \qquad (12.2)$$

be drawn and compared with Fig. 25*a*, it will be seen that the only effect of the third component is to shift the waveform in a vertical

* The symbol $\sim$ indicates that the difference is to be taken without sign; thus if $F_1 = 11/2\pi$, $F_2 = 6/2\pi$, and $K = 2$, $F_1 - KF_2 = -1/2\pi$, but $F_1 \sim KF_2 = 1/2\pi$.

direction: at the parts A, where the third component has the value 5, the waveform is shifted upwards a distance representing 5 units, whereas at the parts B, where the third component has the value -5, the waveform is shifted downwards an equal distance; the shift at other parts of the cycle is proportional to the value of the third component in a similar manner.

In Fig. 25*b* the envelope of Fig. 25*a* is drawn, together with the third component (broken). It is clear that the width of the envelope strip at any part of the wave is the same as that of the function (12.2) at the corresponding place, since both top and bottom envelopes are displaced vertically by the same distance. The normal periodic variation in the strip-width is therefore undisturbed by the addition of the third component. Furthermore, since the envelope of (12.2) has approximately the form

$$4 \cos 2x \text{ (top) or } -4 \cos 2x \text{ (bottom)},$$

(see Section 6), the envelope of (12.1) has approximately the form

$$5 \sin x + 4 \cos 2x \text{ (top) or } 5 \sin x - 4 \cos 2x \text{ (bottom)},$$

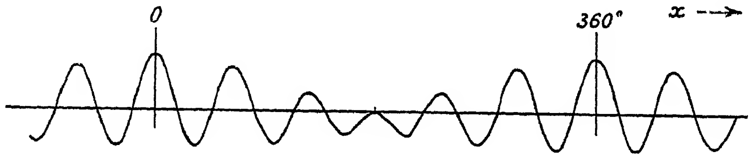


FIG. 26.—Three-component wave, with one frequency equal to the beat frequency of the other two components.

so that the top envelope has a form intermediate between Fig. 22*b* and 22*e*, while the mirror-image, in the centre-line, of the bottom envelope has this form. It may be observed that the *envelope mean*, or wave drawn midway between the top and bottom envelopes, represents the added low-frequency component, for the mean of the two approximate expressions for the envelopes, just given, is $5 \sin x$. This property will be found to be a general one, relating to any wave in which one or more components are of much lower frequency than the remaining components, and is particularly useful in the envelope method of analysis (Chapter IV).

Since a beat envelope is approximately sinusoidal, it is of interest to inquire the effect of adding a third component whose frequency equals the beat frequency. An example of such an addition is given in Fig. 26, which represents the function

$$y = \cos x + 2 \cos 5x + 3 \cos 6x. \quad . \quad . \quad (12.3)$$

Fig. 11 shows the form of the function

$$y = 2 \cos 5x + 3 \cos 6x, \quad . \quad . \quad . \quad (12.4)$$

i.e. the addition of the two higher frequency components of (12.3), and a comparison of the two diagrams shows that the bottom envelope is practically straightened out, and the amplitude of the top envelope is doubled, by the addition of the third component. If this third component were $-\cos x$ instead of $\cos x$, the effect would be to straighten out the top envelope and double the amplitude of the bottom envelope, since the top envelope of (12.4) is approximately $\cos x$ and the bottom envelope $-\cos x$. The fact that in each case one of the envelopes becomes *almost*, and not *exactly*, a straight line, is a result of the variable phase-angle effect discussed in Section 6. In the example described, the amplitude of the added component is the same as the amplitude of the beat envelope; examples of waves of the same type, but with other amplitude ratios, are given in Fig. 19, Chapter IV (p. 103).

The effect of adding a third component of much lower frequency to an existing two-component waveform may be adequately

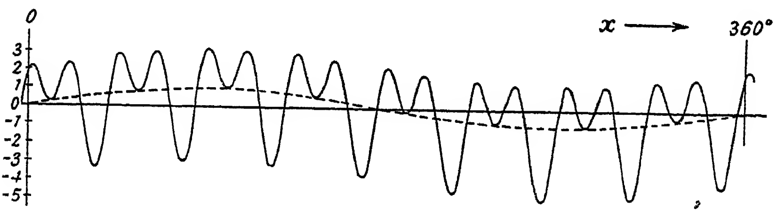


FIG. 27.—Three component wave: low-frequency surge added to wave of Fig. 22*b*, p. 32.

summarised in the statement that the envelope of the resultant three-component wave is obtained by adding the third component to the envelope of the two-component wave.

Fig. 27 illustrates another example of this type of three-component waveform. The diagram represents the function

$$y = \cos x + 2 \sin 8x + 2 \cos 16x. \quad . \quad . \quad (12.5)$$

If the lowest frequency component is omitted from this function the wave has the form of Fig. 22*b*, and again it is seen that the effect of adding the third component is to produce a surge of the waveform.

(ii) *Third component of much higher frequency.*

Fig. 28 shows the effect of adding the third component $8 \cos 12x$ to the two-component waveform $5 \sin x + 10 \sin 2x$. This two-component wave has the form of Fig. 22*g*, and the addition of the third component merely introduces a corresponding high-frequency

"ripple" as in Section 10. The envelopes of the resultant wave are parallel to the two-component wave, which is shown by a broken line in the diagram. This type of three-component wave needs no further comment here.

13. Method of synthesis; tables for synthesis.

In this chapter some representative examples of two-component and three-component waveforms have been examined, since a knowledge of the effects of adding sinusoidal waves together is of inestimable value in the inverse process of analysing a complex wave into its sinusoidal components. A thorough knowledge of these effects, which may conveniently be termed *synthetic* effects, can only be gained as a result of much practical experience in the process of synthesis, and for this reason the reader is strongly recommended to experiment for himself in this direction. Time spent thus is time well spent.

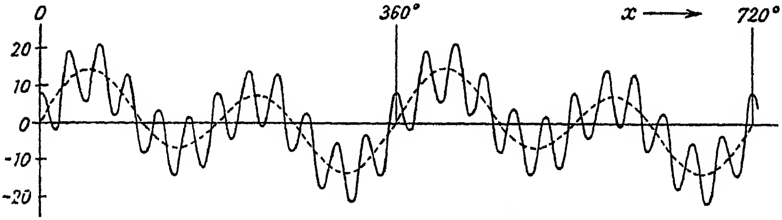


FIG. 28.—Three component wave with high-frequency ripple.

The process of synthesis is facilitated by the use of a table of the values assumed by various sinusoidal functions such as $\sin x$, $\cos x$, $\sin 3x$, $\cos 7x$, etc. The values of $\sin kx$ and $\cos kx$, where k has any integral value up to and including 16, are listed in Table II of Appendix V (p. 258), for values of x at intervals of $3\frac{3}{4}°$ from 0 to 360°. The interval $3\frac{3}{4}°$ has been chosen as it is 1/96 of the total range of 360°; as will be seen in Chapter VII it is useful to have the functions tabulated for 48 equally spaced values of x, and the intermediate values give greater accuracy to synthesised waveforms. Access to a calculating machine, particularly if it be of the electrical automatic type, considerably relieves the tedium of numerical work. In any case, the numerical method of synthesis (by the addition of corresponding values of the component waves) is considerably more accurate than the graphical method, wherein corresponding ordinates are added by means of dividers. The *accurate* graphical synthesis of a waveform having more than three

components requires a high degree of skill and very much more time than does the numerical method.

Two points in connection with numerical waveform synthesis should be noted.

(i) Components of the form $r \sin (kx + \phi)$ must first be expressed in the form $a \cos kx + b \sin kx$, where $a = r \sin \phi$ and $b = r \cos \phi$ (see Section 5).

(ii) When a calculating machine is used, much trouble may be saved by first adding to the components a constant term of sufficient magnitude to ensure that the resultant function is positive over the whole cycle; a constant term equal to the sum of all the amplitudes of the components is sufficient to prevent any of the ordinates of the resultant function becoming negative. Users of calculating machines will readily appreciate this point.

14. Examples for practice.

In order that the reader may test his grasp of the general principles described in the foregoing sections, a number of exercises is given below; successful solution of these will give useful practice and also promote a sense of confidence. It is advised that the waveforms be drawn *after* problems have been solved.

(The answers to the exercises are given on p. 266.)

Determine the constants R and ϕ if $y_1 + y_2 = \mathrm{R} \sin (x + \phi)$ and

(1)	$y_1 = 3{\cdot}2 \sin x$	$y_2 = \cos x$.
(2)	$5 \sin (x + 56°)$	$4 \sin (x - 78°)$.
(3)	$3 \sin x$	$4 \cos x$
(4)	$11{\cdot}5 \sin (x + 30°)$	$- 8 \cos (x - 60°)$.
(5)	$76 \cos (x - 8°)$	$112 \sin (x + 53°)$.

Determine the beat frequency and envelope amplitude in the following waves, and state which is the greater of $\mathrm{S_B}$ and $\mathrm{S_W}$ (see Section 7):

(6) $5 \sin 11t - 8 \cos 12t$.
(7) $\sin t + 0{\cdot}8 \sin 1{\cdot}06t$.
(8) $3 \cos 8t + 4 \cos 8{\cdot}5t$.
(9) $9 \sin 6t + 9 \cos 8t$.
(10) $3{\cdot}8 \cos 5{\cdot}4t + 6{\cdot}5 \cos 5{\cdot}6t$.

State the number of beats and number of crests in a cycle of each of the waves (6)-(10).

Which of the following waves exhibit the properties of (*a*) symmetry, (*b*) skew-symmetry, (*c*) alternance ?

(11) $\sin x - 2 \cos 2x$.
(12) $\sin x - 2 \sin 3x$.
(13) $\cos x + \cos 2x$.
(14) $\sin 5x + 2 \sin 7x$.
(15) $\cos x + \cos 3x - 2 \cos 7x$.

Determine the frequency of the apparent low-frequency surge, if any, in the following waves :

(16) $\sin 3t + 2 \sin 18t$.
(17) $2 \sin 4t - \sin 17t$.
(18) $3 \sin 7t - 2 \sin 13{\cdot}6t$.
(19) $\sin 12t + \sin 24t$.
(20) $2 \sin 6t + \sin 12{\cdot}5t$.

CHAPTER II

GENERAL PROPERTIES OF HARMONIC SERIES

1. Introductory.

In this chapter consideration is first given to the correspondence between a waveform and the physical variation it represents. Harmonic series are defined, and the effect of a change in the basic variable (representing the time measurement or other physical reference quantity) upon the phase-angles of the harmonics is discussed.

The properties of symmetry, skew-symmetry, and alternance are treated in greater detail than in Chapter I, and an account is given of the manner in which these properties indicate the harmonic contents of a waveform.

A standard form of expression for the function represented by a waveform is given in the early part of the chapter; adoption of this standard form, which corresponds to a wavelength of 2π radians or 360°, and to which any waveform can be reduced by a suitable change in the basic variable, leads to a great simplification in the work.

2. Time-base conversion.

A waveform is the graphical representation of a periodic variation in some quantity, the variation being detected and recorded by some mechanical, electrical, or optical device. In a recording system of sound design the basic variable of the waveform (i.e. the variable whose extension is in the horizontal direction in all the examples illustrating Chapter I) is directly proportional to the time-base of the physical variation. The constant of proportionality is termed the *time-base factor* and is found by dividing the length of a portion of the wave by the time-interval represented by that portion. Thus if

l = length of wave, measured in millimetres,
t = time-interval represented by l, measured in seconds,

the time-base factor K_t is given by

$$K_t = l/t, \text{ measured in millimetres/second.} \qquad (2.1)$$

In the case of waveforms recorded directly on tape or film strip, this quantity is more commonly termed the *paper speed* or *film speed*.

The length of wave representing a given time-interval t is found by multiplying the interval by the time-base factor :

$$l = \mathrm{K}_t \,.\, t, \qquad . \quad . \quad . \quad . \quad (2.2)$$

and the time-interval represented by a given length l of wave is found by dividing the length by the factor :

$$t = l/\mathrm{K}_t. \qquad . \quad . \quad . \quad . \quad (2.3)$$

If the period of the physical variation is T seconds, so that the variation $f(t)$ satisfies the condition

$$f(t + k\mathrm{T}) = f(t), \qquad . \quad . \quad . \quad (2.4)$$

where k is any integer (i.e. the variation is exactly repetitive after successive intervals of time T, and T is the smallest interval that can be found to satisfy (2.4)) the wavelength λ of the waveform is found from (2.2) by replacing l and t in that formula by λ and T :

$$\lambda = \mathrm{K}_t \,.\, \mathrm{T}, \qquad . \quad . \quad . \quad . \quad (2.5)$$

and if the waveform is expressed in the form $y = f(x_1)$ then, k being any integer,

$$f(x_1 + k\lambda) = f(x_1). \qquad . \quad . \quad . \quad . \quad (2.6)$$

A great simplification in analysis is achieved by converting the basic variable of a waveform in such a manner that the cycle of the converted wave extends over the range 0-2π. Denoting the new basic variable by x, the wavelength of $f(x)$ is then 2π, so that

$$f(x + 2k\pi) = f(x), \qquad . \quad . \quad . \quad (2.7)$$

where k is any integer. The relation between x and x_1 is given by a comparison of equations (2.6) and (2.7), thus :

$$x_1/\lambda = x/2\pi. \qquad . \quad . \quad . \quad . \quad (2.8)$$

The practical advantage of this conversion lies in the fact that all the components of the waveform can then be expressed in the form $\mathrm{A}_k \sin\,(kx + \phi_k)$, where k is an integer, instead of in the more clumsy form $\mathrm{A}_k \sin\,(k2\pi x/\lambda + \phi_k)$.

3. Harmonic series.

The majority of periodic variations occurring in physical problems give rise to waveforms whose components are members of a simple harmonic series ; i.e. their frequencies are fairly low integral multiples of the frequency of the variation, commonly called the *fundamental frequency*. Thus the components of the waveform have wavelengths of the form

$$\lambda_k = \lambda/k, \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.1)$$

where λ_k = wavelength of component at frequency $k \times$ fundamental frequency,

$\lambda = K_t \, . \, T$,

T = period of variation,

k = an integer,

and k is normally of fairly small magnitude. It is not possible to state a definite upper limit to k, since this depends upon the nature and origin of the variation under consideration. In mechanical vibration analysis values of k not greater than 24 are normally considered, since the components at higher frequencies are usually negligible; in particular cases, however, there may be components at frequencies greater than $24 \times$ fundamental frequency—e.g. arising from gear-tooth impulses. In electrical analysis a much lower limit is frequently imposed on the values of k considered.

As a result of Fourier's Theorem (see Chapter VI) the relation (3.1) holds for *any* periodic variation whatsoever, but in some cases very large values require to be given to k. A particular example of this is found in the vibration analysis of geared aircraft power plants, where the ratio of the rotational speeds of the engine and propeller is frequently not reducible to a simple fraction: two harmonic series of waves are encountered, one having components whose frequencies are multiples of the rotational speed of the propeller, the other having components whose frequencies are multiples of the rotational speed of the camshaft in a 4-stroke engine. The period of the resulting complex vibration is the smallest time which is simultaneously an integral multiple of the time of one propeller revolution and an even multiple of the time of one engine revolution, since the firing cycle in the engine extends over one camshaft or two crankshafts revolutions. For example, suppose that the gear-ratio of the engine-propeller combination is 4/9, i.e. the engine crankshaft revolves nine times in the time-interval taken by the propeller to revolve four times. A cycle of the resulting complex waveform representing, say, the torsional vibrations must include an even number of engine revolutions and a whole number of propeller revolutions; in this case the cycle extends over 8 propeller revolutions and 18 engine revolutions. An engine-excited component at a frequency $2\frac{1}{2} \times$ crankshaft R.P.M. (i.e. $5 \times$ camshaft R.P.M.) thus has a frequency $45 \times$ fundamental frequency of the waveform, while a propeller-excited component at $3 \times$ propeller R.P.M. has a frequency $24 \times$ fundamental frequency of the waveform. If the gears have 16 and 36 teeth respectively, the gear-tooth ripple if present will be at a frequency $36 \times$ propeller R.P.M., and this frequency is $288 \times$ fundamental frequency of the waveform.

It is convenient to employ a simple expression to denote a component sinusoidal wave whose frequency is $k \times$ fundamental frequency, and for this purpose the term k^{th} *harmonic* is used. No matter what the time-base factor K_t may be, the k^{th} harmonic of the waveform represents the k^{th} harmonic of the physical variation. The first harmonic, whose cycle coincides with a cycle of the variation, is termed the *fundamental;* if F is the fundamental frequency of the variation (equal to the reciprocal 1/T of the period T of the variation), and λ the fundamental wavelength of the waveform, the frequency F_k and wavelength λ_k of the k^{th} harmonic are given by

$$F_k = kF, \quad \lambda_k = \lambda/k. \qquad (3.2)$$

The corresponding conversion factors are listed in Table I for harmonics up to, and including, the 12th.

TABLE I

Frequencies and wavelengths of first 12 harmonics

F = fundamental frequency

λ = fundamental wavelength

k.	F_k/F.	λ_k/λ
1	1	1·000
2	2	0·500
3	3	0·333
4	4	0·250
5	5	0·200
6	6	0·167
7	7	0·143
8	8	0·125
9	9	0·111
10	10	0·100
11	11	0·091
12	12	0·083

The physical variation $f(t)$ may be expressed in the form *

$$\begin{aligned} f(t) &= R_0 + \sum_k R_k \sin(2\pi F_k t + \phi_k) \\ &= R_0 + R_1 \sin(2\pi F_1 t + \phi_1) + R_2 \sin(2\pi F_2 t + \phi_2) \\ &\qquad + R_3 \sin(2\pi F_3 t + \phi_3) + \text{etc.} \\ &= R_0 + R_1 \sin(\omega t + \phi_1) + R_2 \sin(2\omega t + \phi_2) \\ &\qquad + R_3 \sin(3\omega t + \phi_3) + \text{etc.} \end{aligned} \qquad (3.3)$$

where $\omega = 2\pi F$

* Wherever the notation $\sum\limits_k$ is used, it is to be understood that the expression which follows is to be summed for all positive integral values of k.

and the waveform can be expressed in the form

$$
\begin{aligned}
y &= A_0 + \sum_k A_k \sin (kx + \phi_l) \\
&= A_0 + A_1 \sin (x + \phi_1) + A_2 \sin (2x + \phi_2) \\
&\qquad\qquad + A_3 \sin (3x + \phi_3) + \text{etc.} \qquad . \quad (3.4)
\end{aligned}
$$

where $x = \dfrac{2\pi}{\lambda} x_1,$

$\lambda = K_t \,.\, T = K_t/F,$

$x_1 = K_t \,.\, t,$

i.e. $x = \dfrac{2\pi F}{K_t} K_t \,.\, t = 2\pi Ft = \omega t.$

The constants A_1, A_2, etc., of (3.4) are, in a recording system free from distortion (see Chapter IX, p. 225), proportional to the constants R_1, R_2, etc., of (3.3). The increment δy of the wave-function y corresponding to an increment $\delta f(t)$ in the physical quantity $f(t)$ is given by

$$\delta y = M \,.\, \delta f(t), \qquad . \quad . \quad . \quad (3.5)$$

where M is the *calibration constant* or *magnification* of the recording system, and the relation between A_k and R_k is of the same form :

$$A_k = M \,.\, R_k. \qquad . \quad . \quad . \quad . \quad (3.6)$$

The constants A_0 and R_0 may or may not be connected by the relation (3.6), dependent upon the nature of the recording system. In any case, these quantities refer to a particular choice of arbitrary datum levels. Since the average value of sin kx or cos kx over the range 0-2π is zero, if k is an integer, the constant A_0 is the average value of the height of the waveform y (3.4) above this datum level.

4. Phase-angle : change of basic variable.

The form (3.4) of expression for a waveform will be adopted as a standard form in the remainder of this work. Utilising the result (5.1) from Chapter I, the standard form

$$y = A_0 + \sum_k A_k \sin (kx + \phi_k), \qquad . \quad . \quad (4.1)$$

can be rewritten as

$$
\left.
\begin{aligned}
y &= A_0 + \sum_k a_k \cos kx + \sum_k b_k \sin kx \\
\text{where} \quad a_k &= A_k \sin \phi_k \\
b_k &= A_k \cos \phi_k
\end{aligned}
\right\} \qquad . \quad (4.2)
$$

Both forms of expression (4.1) and (4.2) are convenient on different occasions. The former indicates directly the severity of any

harmonic, whereas the latter is sometimes more convenient when a change of basic variable is considered, as will now be discussed.

Let the basic variable x be changed to x_1, where

$$x_1 = x - s, \quad (x = x_1 + s). \qquad (4.3)$$

This substitution is equivalent to shifting the reference point, for measurement along the x-axis, to the point $x = s$. The k^{th} harmonic becomes $A_k \sin [kx_1 + (\phi_k + ks)]$ and (4.1) becomes

$$y = A_0 + \sum_k A_k \sin [kx_1 + (\phi_k + ks)]. \qquad (4.4)$$

The phase-angle ϕ_k of the k^{th} harmonic is increased by the amount ks, which is clearly proportional to the number k.

In alternating-current electrical design it is sometimes essential to consider the effect of the circuit on the phase-angles of the components of the alternating current. If this effect is to increase all the phase-angles (or decrease all of them) by amounts proportional to the harmonic number k, and there is no distortion of amplitudes, it is evident that the output waveform is the same shape as the input waveform, with a general shift in the time-base. For example, suppose the input wave to be of the form (4.1), while the output wave is of the form

$$y' = B_0 + \sum_k B_k \sin (kx + \phi_k + \psi_k), \qquad (4.5)$$

where the B_k are proportional to the A_k of (4.1), the constant of proportionality being the *overall magnification* M of the circuit (which is independent of k) and the ψ_k are proportional to k ($\psi_k = k\psi$); then (4.5) can be written as

$$y' - B_0 = M \sum_k A_k \sin [k(x + \psi) + \phi_k],$$

and a comparison of this last equation with (4.1) shows that the sole effect of such phase-distortion is to shift the waveform bodily along the axis of the basic variable x.

Example.

Let $$y = A_0 + A_1 \sin (x + 10°) + A_2 \sin (2x + 10°) + A_3 \sin (3x + 15°) + A_7 \sin (7x + 20°),$$

and let $x_1 = x + 5°$. Plotting the waveform against x_1 instead of against x will then result in a shift of the whole wave along the x-axis, through a distance representing 5° in the positive direction, and

$$y = A_0 + A_1 \sin (x_1 + 5°) + A_2 \sin 2x_1 + A_3 \sin 3x_1 + A_7 \sin (7x_1 - 15°).$$

5. Change of basic variable by multiples of a quarter-period.

Particular cases of the general change of basic variable are those wherein the wave is shifted through a distance which represents one, two or three right-angles (or in the physical variation (3.3), through one, two or three quarter-periods).

(i) *One right-angle (quarter-period).*

In the standard form (4.2), consider the effect of changing the basic variable to x_1, where $x_1 = x + 90°$. This is equivalent to shifting the wave through a distance representing a right-angle, or a quarter wave-length, in the positive direction. The k^{th} harmonic terms are then

$$a_k \cos k(x_1 - 90°) + b_k \sin k(x_1 - 90°). \quad . \quad . \qquad (5.1)$$

This expression can be simplified by utilising the results (2.5) from Chapter I. The final form depends upon the remainder when k is divided by 4. N being an integer, if $k = 4\text{N}$ the harmonic terms (5.1), which can be rewritten

$$a_k \cos (kx_1 - 90k°) + b_k \sin (kx_1 - 90k°), \quad . \qquad (5.2)$$

are of the form

$$a_k \cos kx_1 + b_k \sin kx_1, \quad . \quad . \quad . \qquad (5.3)$$

since the phase-angle in both terms is a multiple of 360°. Thus the 4th, 8th, 12th, etc., harmonic terms of (4.2) are unaltered by the change of variable.

If $k = 4\text{N} + 1$, then by reason of the fourth equation of (2.5) in Chapter I the terms (5.2) can be written as

$$\begin{aligned} a_k \cos (kx_1 - 90°) &+ b_k \sin (kx_1 - 90°) \\ &= a_k \sin kx_1 - b_k \cos kx_1 \end{aligned} \qquad (5.4)$$

Thus the 1st, 5th, 9th, etc., cosine harmonic terms become sine terms, while the corresponding sine harmonic terms become cosine terms with a change of sign.

If $k = 4\text{N} + 2$ (i.e. k is even), the terms (5.2) can be written as

$$\begin{aligned} a_k \cos (kx_1 - 180°) &+ b_k \sin (kx_1 - 180°) \\ &= -(a_k \cos kx_1 + b_k \sin kx_1) \end{aligned} \qquad (5.5)$$

The remaining even harmonics therefore have their signs reversed by a quarter-period shift.

If $k = 4\text{N} + 3$, the terms (5.2) can be written as

$$\begin{aligned} a_k \cos (kx_1 + 90°) &+ b_k \sin (kx_1 + 90°) \\ &= -a_k \sin kx_1 + b_k \cos kx_1, \end{aligned} \qquad (5.6)$$

and the effect of the quarter-period shift on the 3rd, 7th, 11th, etc., harmonics is to change cosine terms to sine terms with a change of sign, and sine terms to cosine terms.

The effects of this phase-shift on the various components are illustrated in Fig. 1. At (*a*), the full line represents $\sin \theta$, and the broken line $\cos \theta$, where $\theta = kx$. At (*b*), the origin ($\theta = 0$) has been displaced through a distance representing 90° to the left, so that θ now is equal to $(kx + 90°)$; if k is of the form $4N + 1$, the full line now represents $-\cos \theta$ and the broken line $\sin \theta$. Thus if k is of the form $4N + 1$, the diagram illustrates the conversions given in formula (5.4). In a similar manner, (*c*) and (*d*) illustrate the conversions given in formulæ (5.5) and (5 6).

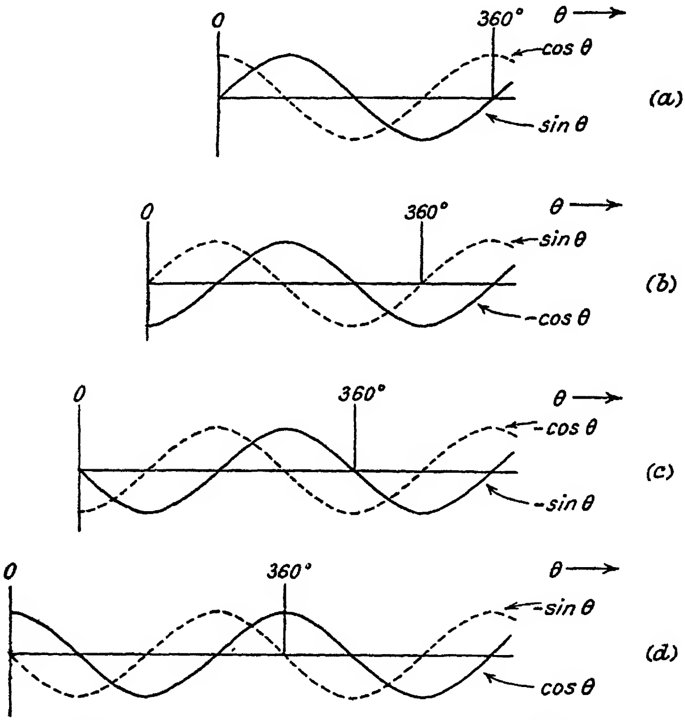


FIG. 1.—Effects of phase-shift through a multiple of a quarter-period.

These results (5.3) to (5.6) are summarised in Table II, wherein C stands for "cosine" and S for "sine." The entries in the table show to what the sine and cosine terms of the first 12 harmonics are converted by increasing the basic variable by a quarter of the fundamental period.

(ii) *Two right-angles (half-period).*

If the basic variable x in (4.2) be changed to x_2, where $x_2 = x + 180°$, the wave is effectively shifted along the x-axis through a distance

representing two right-angles, or half a wave-length, in the positive direction. The k^{th} harmonic terms are then

$$a_k \cos k(x_2 - 180°) + b_k \sin k(x_2 - 180°),$$

i.e.
$$a_k \cos (kx_2 - 180k°) + b_k \sin (kx_2 - 180k°). \quad . \quad (5.7)$$

If k is even, (5.7) can be written as

$$a_k \cos (kx_2 - 360\mathrm{N}°) + b_k \sin (kx_2 - 360\mathrm{N}°), \quad . \quad (5.8)$$

where N is an integer, and (5.8) is clearly equivalent to (5.3), since addition of a multiple of 360° to the basic variable does not alter the value of a sine or cosine function. The even harmonics are therefore unaffected by a half-period shift.

If k is odd, (5.7) can be written as

$$a_k \cos (kx_2 - 180°) + b_k \sin (kx_2 - 180°)$$
$$= -(a_k \cos kx_2 + b_k \sin kx_2), \quad (5.9)$$

since the sign of a sine or cosine function is reversed when the argument (i.e. the basic variable) is increased by 180°. The odd harmonics therefore have their signs reversed by a half-period shift.

The results (5.8), (5.9) are summarised in Table III.

(iii) *Three right-angles (three-quarter period).*

By exactly the same method as used in (i) above it can easily be shown that the effects of a three-quarter period shift in the positive direction are as listed in Table IV. It is to be noted that Tables II and IV agree for the even harmonics, and show a sign-reversal for the odd harmonics; this result could of course be obtained by considering the three-quarter period shift as made in two stages: first, a shift of a quarter-period, giving the results of Table II, and then a further shift of a half-period, which according to Table III leaves the even harmonics unaffected and reverses the signs of the odd harmonics. Alternatively, the three-quarter period shift may be regarded as a shift of a quarter-period in the negative direction, and the values listed in Table IV read off from those given in Table II by reverse entry: thus since when $k = 1$, S becomes $-$ C and C becomes S in Table II, by reverse entry it is found that in Table IV $-$ C must become S (i.e. C becomes $-$ S) and S becomes C, and similarly for the other harmonics.

Examples.

(i) Let $y = 20 + 5 \cos x + 7 \cos 3x + 3 \sin 4x$.

Replacing x by $x_1 - 270°$, i.e. shifting the origin of measurement along the x-axis through a distance representing 270° in the negative direction, corresponding to a shift of the wave, with respect to the

TABLE II

Effect of increasing the basic variable by a quarter-period

S = sine, C = cosine

$k=$	1	2	3	4	5	6	7	8	9	10	11	12
S terms become :	−C	−S	C	S	−C	−S	C	S	−C	−S	C	S
C terms become :	S	−C	−S	C	S	−C	−S	C	S	−C	−S	C

TABLE III

Effect of increasing the basic variable by a half-period

$k=$	1	2	3	4	5	6	7	8	9	10	11	12
S terms become :	−S	S	−S	S	−S	S	−S	S	−S	S	−S	S
C terms become .	−C	C	−C	C	−C	C	−C	C	−C	C	−C	C

TABLE IV

Effect of increasing the basic variable by a three-quarter period

$k=$	1	2	3	4	5	6	7	8	9	10	11	12
S terms become :	C	−S	−C	S	C	−S	−C	S	C	−S	−C	S
C terms become :	−S	−C	S	C	−S	−C	S	C	−S	−C	S	C

origin, through an equal distance in the positive direction, Table IV gives

$$y = 20 - 5 \sin x_1 + 7 \sin 3x_1 + 3 \sin 4x_1.$$

A simple change of basic variable has reduced the mixed sine-and-cosine series expression for y to one involving sine terms only.

(ii) Consider the series

$$y = A_0 + b_1 \sin x + b_3 \sin 3x + b_5 \sin 5x + \text{etc.},$$

which involves (apart from the constant term A_0) only odd sine terms. By shifting the wave along the x-axis through a distance representing a quarter-period either in the positive direction (Table

II) or in the negative direction (Table IV), these sine terms are changed to cosine terms. Putting $x = x_1 - 90°$, Table II gives

$$y = A_0 - b_1 \cos x_1 + b_3 \cos 3x_1 - b_5 \cos 5x_1 + \text{etc.},$$

while putting $x = x_2 - 270° = x_2 + 90°$, Table IV gives

$$y = A_0 + b_1 \cos x_2 - b_3 \cos 3x_2 + b_5 \cos 5x_2 - \text{etc.}$$

Thus a wave involving only odd sine terms can be represented by a series involving only odd cosine terms, and *vice versa*, by means of a quarter-period shift of the wave along the axis of the basic variable in either direction. The waveform illustrated in Fig. 23*a* of Chapter I (p. 36) is a particular example of this phenomenon, and that in Fig. 23*c* of that chapter is another. These waveforms may be regarded as the sum of first and third sine components or first and third cosine components, according to the choice of origin.

It should be noted that replacing x by $x_1 - a$ is equivalent to shifting the origin of measurement along the x-axis through a distance representing a in the negative direction, corresponding to a shift of the wave, relative to the origin, through an equal distance in the positive sense; for the point $x_1 = 0$ is the point $x = -a$, since $x = x_1 - a$. Similarly, replacing x by $x_1 + a$ is equivalent to shifting the origin of measurement through a distance representing a in the positive sense, corresponding to a shift of the wave, relative to the origin, through an equal distance in the negative direction; for the point $x_1 = 0$ is now the point $x = a$.

6. Symmetry, skew-symmetry and alternance.

The properties of symmetry, skew-symmetry and alternance are of great use in indicating the nature of the components of a waveform. These properties have been mentioned in Chapter I in connection with sinusoidal waves, and will now be discussed at greater length.

Confining attention to waveforms of the type (4.1) or (4.2), i.e. with a wavelength of 2π radians (since any other waveform can be reduced to this type by means of a simple transformation such as is discussed in Section 2), it is immediately clear that **if all the harmonic components of a wave are symmetrical about a certain value X of x, then the wave itself is symmetrical about $x = X$, and conversely.** For the property of symmetry in a function $f(x)$ can be expressed in the form

$$f(X + a) = f(X - a), \qquad . \quad . \quad . \quad (6.1)$$

where a has any value and $x = X$ is the line of symmetry; and if

the functions $f_1(x)$, $f_2(x)$, . . . $f_n(x)$ all satisfy the condition (6.1) their sum does likewise. Denoting this sum by $F(x)$, so that

$$F(x) = f_1(x) + f_2(x) + \ldots \quad \ldots f_n(x),$$

then $F(X + a) = f_1(X - a) + f_2(X + a) + \ldots \ldots f_n(X + a)$,
and $F(X - a) = f_1(X - a) + f_2(X - a) + \ldots \ldots f_n(X - a)$.

Since all the functions on the right-hand side of these two equations satisfy (6.1), it follows that

$$F(X - a) = F(X - a).$$

Thus if all the components of a waveform are symmetrical about the same value $x = X$, the waveform itself is symmetrical about that value. Conversely, let $g(x)$ be a function *not* symmetrical about the value $x = X$, so that

$$g(X + a) \neq g(X - a),$$

and let $G(x)$ denote the function $F(x) + g(x)$. Then

$$\begin{aligned} G(X + a) &= F(X + a) + g(X + a) \\ &= F(X - a) + g(X + a) \\ &\neq F(X - a) + g(X - a) \\ &\neq G(X - a), \end{aligned}$$

so that if any of the harmonic components is not symmetrical about $x = X$ the resultant wave is not. It therefore follows that **all the components of a wave are symmetrical about those values of the basic variable which determine the lines of symmetry, if any, in the resultant wave.**

The same argument applies to skew-symmetry, the condition for which (about the value $x = X$) is

$$f(X + a) = -f(X - a) \quad . \quad . \quad . \quad (6.2)$$

for all values of a.

If the waveform $f(x)$, having a wavelength 2π, is symmetrical or skew-symmetrical about the value $x = X$, it is also symmetrical or skew-symmetrical about the value $x = X + \pi$. This result is easily proved: taking first the symmetrical case,

$$f(X + a) = f(X - a),$$

and if $b = \pi - a$, this equation gives

$$f(X + \pi - b) = f(X - \pi + b).$$

But since 2π is the wavelength of $f(x)$,

$$f(X - \pi + b) = f(X + \pi + b),$$

hence $$f[(X + \pi) + b] = f[(X + \pi) - b],$$

and this is the condition (6.1) for symmetry about the value $x = X + \pi$. Fig. 2*a* illustrates this property; the function depicted is symmetrical about A and B, and if x_A, x_B are the corresponding values of x for these points,

$$x_B - x_A = \pi.$$

Similar reasoning demonstrates the result for the case of skew-symmetry, and Fig. 2*b* shows an example; the function illustrated is skew-symmetrical about A and B, and $x_B - x_A = \pi$.

The interval of alternance of a waveform of the type (4.1) or

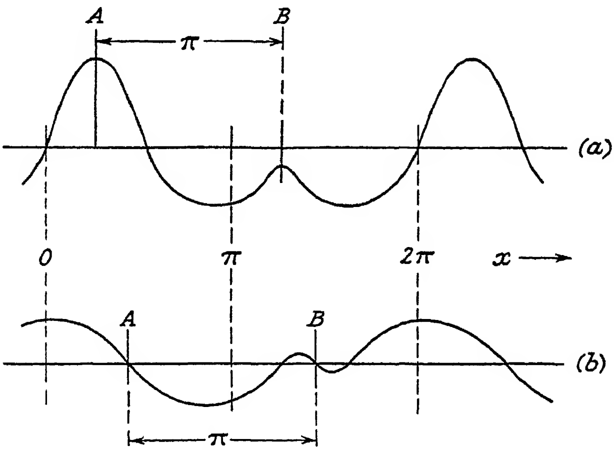


FIG. 2.—Symmetry and skew-symmetry: (*a*) symmetrical periodic wave; (*b*) skew-symmetrical periodic wave.

(4.2) can only be the half-period (π radians or 180°). If the interval of alternance of the waveform $f(x)$ is I_A, so that

$$f(x + I_A) = -f(x), \quad . \quad . \quad . \quad (6.3)$$

a moment's consideration shows that I_A must be of the form π/K, where K is an integer, in order that $f(x)$ may have a wavelength of 2π radians. Furthermore, by a double application of the condition (6.3) it is seen that

$$f(X + 2I_A) = f(x). \quad . \quad . \quad . \quad (6.4)$$

Now if $I_A = \pi$, (6.4) states that the wavelength of $f(x)$ is 2π; but if K is other than unity, a smaller wavelength is indicated. thus if $K = 2$, the wavelength is π or 180°. Such a result indicates that a further change of basic variable, in the manner described in Section 2, is required so that 2π may be the *smallest* value of λ such that $f(x + \lambda) = f(x)$.

Reasoning similar to that used in connection with symmetry and skew-symmetry shows that if a waveform is alternant, all the harmonic components must be alternant with the same interval as the wave; and it has just been demonstrated that the only possible interval of alternance for a wave, whose wavelength is 2π radians, is π radians or half a wavelength. It therefore follows that all the harmonic components of an alternant waveform must be alternant, half the wavelength of the waveform being the interval of their alternance. Now, if k is an even integer,

$$\binom{\sin}{\cos} k(x+\pi) = \binom{\sin}{\cos} kx,$$

while if k is an odd integer,

$$\binom{\sin}{\cos} k(x+\pi) = -\binom{\sin}{\cos} kx.$$

It is therefore apparent that the harmonic series for an alternant waveform includes only odd harmonics.

"*Disguised*" *skew-symmetry and alternance.* The properties of symmetry, skew-symmetry and alternance are intrinsic pro-

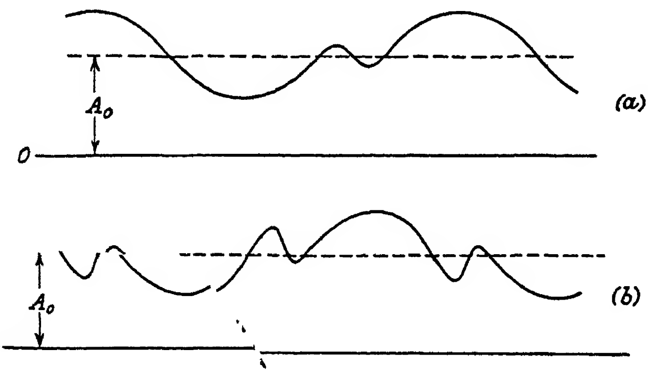


FIG. 3.—"Disguised" skew-symmetry and alternance.

perties of a wave, i.e. they do not depend upon the choice either of the basic variable or of the constant term which determines the height of the centre-line above the arbitrary datum-level. The curve shown in Fig. 3*a* represents the function obtained by adding a constant A_0 to the function represented by Fig. 2*b*. The shape of the wave has not been affected by this addition, but the condition (6.2) no longer holds good. The property possessed by the curve is known as *disguised skew-symmetry,* and the mathematical con-

dition analogous to (6.2) is easy to formulate. Let $y = g(x)$ be the equation to the curve illustrated. Then

$$g(x) = f(x) + A_0,$$

where $y = f(x)$ is the equation to the curve illustrated in Fig. 2*b*. Equation (6.2) then gives

$$g(X + a) - A_0 = -g(X - a) + A_0,$$

i.e.
$$g(X + a) + g(X - a) = 2A_0, \quad . \quad . \quad . \quad . \quad (6.5)$$

for all values of a.

(6.5) is the required mathematical condition for disguised skew-symmetry about the value $x = X$, and states that the added constant A_0 is the mean of the two values $g(X + a)$ and $g(X - a)$ which would be equal in magnitude and opposite in sign if the skew-symmetry were not disguised. When $A_0 = 0$, the equation (6.5) reduces to (6.2).

Similarly, the curve illustrated in Fig. 3*b* possesses a disguised type of alternance; in this case the mathematical condition analogous to (6.3) is

$$g(x) + g(x + \pi) = 2A_0, \quad . \quad . \quad . \quad (6.6)$$

so that A_0 is the mean of the two values $g(x)$ and $g(x + \pi)$ which would be equal in magnitude and opposite in sign if the alternance were not disguised. (6.6) of course reduces to (6.3) if $A_0 = 0$.

The same symbol A_0 has deliberately been used in this connection as was employed in the harmonic series (4.1) and (4.2), for if a wave is skew-symmetrical or alternant there is no constant term in the harmonic series representing the waveform, and if the skew-symmetry or alternance is disguised the constant term in the series is equal to the distance by which the wave is displaced vertically from the position in which the skew-symmetry or alternance is not disguised. This result is easily found by considering the average value of the function over a cycle ($b \leqslant x \leqslant b + 2\pi$, where b has any value). Since the average value of a sine or cosine function, taken over an integral number of cycles, is zero, the average value of the function is the constant term A_0 in the series (4.1) or (4.2); and taking first the case of disguised skew-symmetry, if D represents temporarily the distance by which the wave is displaced vertically from the "undisguised" position, (6.5) gives

$$g(X + a) + g(X - a) = 2D.$$

The average value of $g(x)$ over a cycle is found by performing the integration

$$\frac{1}{2\pi}\int_0^{\pi}[g(X + a) + g(X - a)]da,$$

which reduces to D by reason of the preceding equation. Thus $D = A_0$.

Similarly, in the case of disguised alternance (6.6) gives

$$g(x) + g(x + \pi) = 2D,$$

and the average value of $g(x)$ over a cycle can be found by performing the integration

$$\frac{1}{2\pi}\int_0^\pi [g(x) + g(x + \pi)]dx,$$

which reduces to D; thus in this case also $D = A_0$.

7. Determination of harmonic contents by considerations of symmetry, etc.

By a simple change in the basic variable x, a waveform which is symmetrical or skew-symmetrical about the values $x = X$, $x = X + \pi$, can be converted into one for which the critical values are $x = 0$, $x = \pi$. Assuming that such a transformation has been made, so that the centres of symmetry, etc., are at the beginning and middle of a cycle, the following two rules can be formulated:

(i) the harmonic series (4.2) representing a symmetrical waveform contains no sine terms, and

(ii) the harmonic series (4.2) representing a skew-symmetrical wave contains no cosine term.

The rules follow immediately from the fact that only cosine terms are symmetrical about the values $x = 0$ and $x = \pi$, and only sine terms are skew-symmetrical about these values. A third rule has already been found in Section 6 for the case of alternance:

(iii) the harmonic series (4.1) or (4.2) representing an alternant waveform contains no even harmonics.

In these rules it is of no consequence whether the alternance or skew-symmetry is disguised or not, since this affects only the constant term in the series.

Obsolescent terms for symmetrical and skew-symmetrical waves are "even" and "odd" waves. It is fortunate that these terms are going out of fashion, as they can be very misleading. It appears to be much more logical to reserve the use of the terms "odd" and "even" to denote the nature of the harmonic numbers of the components, i.e. to use them in their ordinary arithmetical sense.

The same waveform can be both symmetrical and skew-symmetrical, although naturally the critical values will be different

for the two properties. Consideration shows that if a wave is symmetrical about the values $x = 0$ and $x = 180°$, the only values about which it can be skew-symmetrical are 90° and 270°; similarly, if a wave is skew-symmetrical about $x = 0$ and $x = 180°$, the only values about which it can simultaneously be symmetrical are 90° and 270°. Now, all sine terms in (4.2) are skew-symmetrical about $x = 0$ and $x = 180°$, and all cosine terms in (4.2) are symmetrical about the same values; furthermore, even cosine and even sine terms are respectively symmetrical and skew-symmetrical about the values $x = 90°$ and $x = 270°$, while for odd harmonics the properties are interchanged, odd cosine terms being skew-symmetrical and odd sine terms symmetrical, about these values. These results are summarised in Table V, together with an indication of which terms are alternant with the interval 180°.

TABLE V

Symmetry, skew-symmetry, and alternance

sym. = symmetrical; skew = skew-symmetrical;
alt. = alternant with the interval 180°

Terms.	Odd sine.	Even sine.	Odd cosine.	Even cosine.
Critical values:				
0°	skew	skew	sym.	sym.
90°	sym.	skew	skew	sym.
180°	skew	skew	sym.	sym.
270°	sym.	skew	skew	sym.
	alt.	—	alt.	—

A wave which possesses any two of the three properties possesses the third also, as examination of Table V shows. This result can also be obtained by a consideration of Tables II and III. Suppose, for example, a wave is alternant and is also skew-symmetrical about the values $x = 0$ and $x = 180°$. Table III shows that since alternance implies that an increase of half a period in the basic variable produces a change of sign, the series representing an alternant waveform contains odd harmonics only. Also, the wave in its original state is represented by a series containing only sine terms, and Table II shows that an increase of a quarter-period in the basic variable changes all odd sine terms to cosine terms. The original wave is therefore symmetrical about $x = 90°$ and $x = 270°$.

It is evident that by examination of a waveform for properties

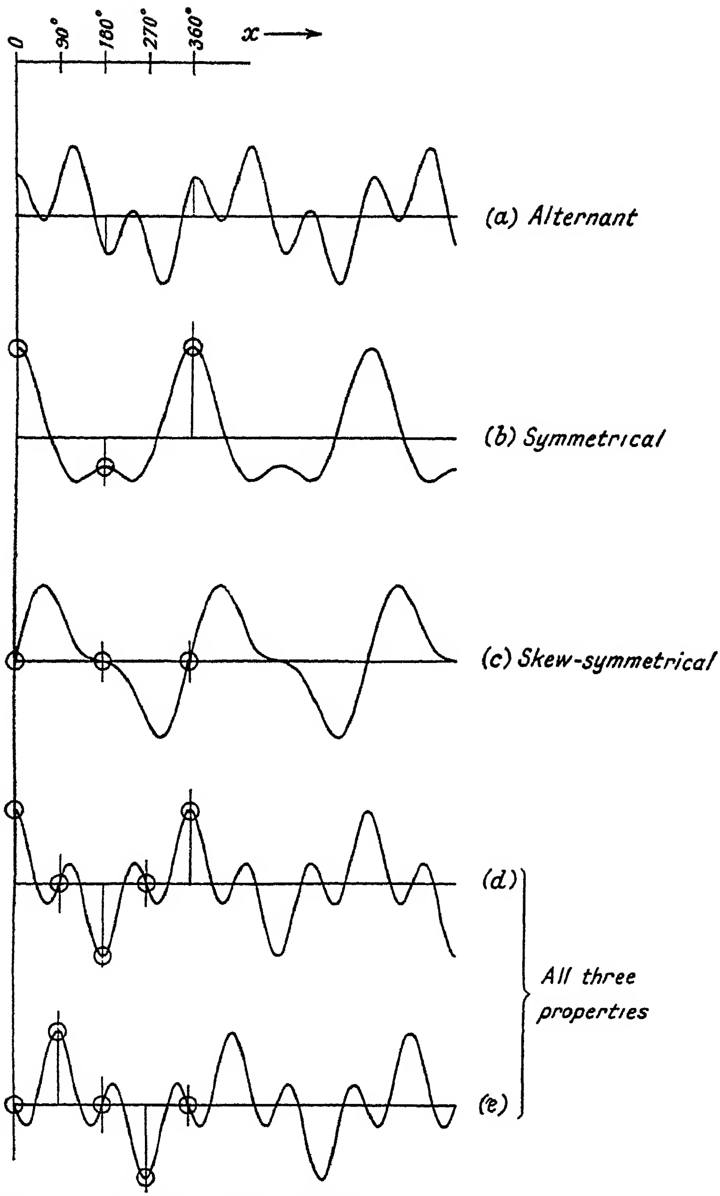


FIG. 4.—Examples of waveforms showing properties of symmetry, etc.

of symmetry, skew-symmetry and alternance, much information can be gained as to the harmonic contents. Thus, for example, a wave which is symmetrical can be represented by a series of cosine

terms, and one which is both alternant and symmetrical can be represented by a series either of odd sine terms only or of odd cosine terms only. Assuming that the point $x = 0$ is chosen so as to bring the critical points, if any, to one or more of the values $x = 0$, 90°, 180°, 270°, Table VI summarises the analytical results.

Fig. 4 shows examples of these types of waveform. The functions represented are as under (letter references correspond to those in Table VI):

$$(a)\ y = \sin x + \cos 3x.$$
$$(b)\ y = 2 \cos x + \cos 2x.$$
$$(c)\ y = 2 \sin x + \sin 2x.$$
$$(d)\ y = \cos x + \cos 3x.$$
$$(e)\ y = \sin x - \sin 3x.$$

Table VI

Harmonic contents determined by consideration of symmetry, etc.

Critical values (i.e. centres of symmetry or skew-symmetry) given in parentheses.

Ref	Type of wave-form	Harmonic terms
a	Alternant	odd sine and odd cosine
b	Symmetrical (0 and 180°)	cosine
c	Skew-symmetrical (0 and 180°)	sine
d	Alternant and symmetrical (0 and 180°), hence also skew-symmetrical (90° and 270°)	odd cosine
e	Alternant and skew-symmetrical (0 and 180°), hence also symmetrical (90° and 270°)	odd sine

8. Examples for practice.

Some exercises on the work covered in the chapter are given below, in order that the reader may satisfy himself that he has grasped the principles. Answers will be found at the end of the book (p. 266).

(1) The wavelength of a calibration record, representing the 50 C.P.S. electric mains voltage variation, is 15 mm. Find the length of a record covering a time of one minute, and also the wavelength of a recorded wave whose frequency is 2600 C.P.M. A component has a wavelength of 8·65 mm.; what is its frequency?

(2) The fundamental wavelength of a complex record is 6 in. Find the wavelengths of the 3rd, 6th, and 12th harmonics. Which harmonics have wavelengths of 0·75 and 0·25 in.?

(3) An aircraft engine-propeller combination has a gear-ratio 6 : 13 (i.e. 6 revolutions of the propeller occur in the same time as 13 revolutions of the engine). If harmonics from both sources are present in the recorded waveform, and the film speed is such that 1·8 in. of record corresponds to one engine revolution, find the length of record corresponding to a cycle of the waveform, if (*a*) some of the half-order engine harmonics (e.g. $1\frac{1}{2} \times$ engine crankshaft rotational speed) are present, and (*b*) only integral harmonics are present in the engine series. Find the harmonic numbers, with reference to the complex waveform, of the 3rd engine and 4th propeller harmonics in both cases (*a*) and (*b*).

(4) Repeat exercise 3 for the case where the gear-ratio is 1 : 2.

(5) Reduce the following harmonic series to a form in which the fundamental component is a sine term without phase-angle:

$$3{\cdot}2 + 1{\cdot}5 \sin (x - 19^\circ) + 2{\cdot}7 \sin (2x + 7^\circ) + 0{\cdot}8 \sin (4x - 85^\circ)$$

(6) Reduce the following harmonic series to a form in which the third harmonic is a cosine term without phase-angle:

$$\sin x + 2{\cdot}5 \sin (2x - 72^\circ) + \sin (3x + 45^\circ) + 3{\cdot}6 \sin 4x.$$

(7) Find the angle between the rotating vectors representing the pairs of waves:

(*a*) $\sin (2x - 10^\circ)$, $\sin (2x + 75^\circ)$.
(*b*) $\sin 3(x - 15^\circ)$, $\sin 3(x - 45^\circ)$.
(*c*) $\sin (4x - 12^\circ)$, $\sin 4(x + 52^\circ)$.
(*d*) $\sin (5x - 90^\circ)$, $\sin 5(x - 18^\circ)$.

(8) State which of the following waveforms are (*a*) alternant, (*b*) symmetrical, (*c*) skew-symmetrical. The numbers are harmonic numbers, and C and S respectively stand for *cosine* and *sine* terms; thus the first example contains 1st, 3rd, and 7th sine and cosine terms.

(i)	1S,	1C,	3S,	3C,	7S,	7C.
(ii)	1S,	2C,	3S,	5S,	6C.	
(iii)	1C,	5C,	9C.			
(iv)	1C,	3C,	7S,	11C.		
(v)	3C,	4C,	6C.			
(vi)	2S,	3S,	6S.			

(9) By consideration of the property of alternance, deduce a method of separating the odd and even harmonics in a waveform. (See Chapter V, p. 121.)

(10) By consideration of the properties of symmetry and skew-symmetry deduce a method of separating the sine and cosine terms in a waveform. (See Chapter V. p. 123.)

CHAPTER III

BASIC ANALYSIS OF RECORDED WAVEFORMS

1. Introductory.

The preceding chapters have been concerned with the properties of sine-waves in combination and of harmonic series. The process of analysis, whereby the harmonic contents of a given periodic variation can be determined, is described in the succeeding chapters ; first, however, various points in connection with basic measurements of recorded waveforms must be discussed.

In this chapter are considered those waves of which the main characteristics can be observed directly in the waveform. A pure sine-wave, for example, requires no analysis into simpler components ; yet the information that can be extracted from a record of this type is not so extracted until several measurements have been made. The position of all the crests and all the troughs is clear, yet the amplitude must be *measured* and the frequency found from further *measurements.*

Analysis by the envelope method, as described in Chapter IV, breaks down the complex wave into its sine-wave components by a systematic scheme of reduction. These sine components are not usually directly evident in the waveform, but the scheme of reduction enables the analyst to determine where to make his measurements, so as to obtain all the information (concerning amplitudes, frequencies and phase-angles) he desires. The present chapter describes in detail the manner in which these measurements should be made, and the method of converting them to the required results.

Frequency determination is first discussed, consideration being given both to absolute frequency values and to "order numbers," which express the relationship between the absolute frequency of a wave and some standard frequency, such as the rotational speed in R.P.M. of an engine. Measurement of amplitudes is then described, with a passing reference to the overall amplitude of very complex waves, whose cycles may extend over a considerable length of the record. The chapter concludes with a description of the method of determining phase-angles.

2. Frequency determination.

One of the prime objects of the analysis of waveforms is to determine the frequencies of the constituent components from

which the wave has been built up. In this present section is considered the determination of those frequencies which can be "spotted" in a waveform; the waves illustrated in Figs. 18 to 28 in Chapter I are cases where two or more frequencies can actually be seen, and it is with such cases as this that the following paragraphs are concerned. The methods described are fundamental and can be extended, as required, to deal with more complicated waveforms.

It is essential to possess information concerning the time-base of the record (see Chapter II, Section 2). The practical details of the methods whereby this information may be recorded are matters of instrumentation and recording technique, and as such do not come within the scope of this book; the outstanding requirements in this connection are listed in Chapter IX, and for the present it will be assumed that the waveform incorporates some indication of the independent variable, or time-base.

Note.—It is convenient to refer to the independent variable as a time-base, even when the recording mechanism is so arranged that the variable represented by the vertical excursion of the trace is plotted as a function of some variable other than time—linear or angular displacement of some part of a machine, for example.

It may be desired to determine the frequency directly, expressing the result as a number of cycles per second or per minute, or to refer it to some other standard; and in some cases both methods may be used simultaneously. The standard most frequently employed in the analysis of vibrations occurring in structures when the disturbances are caused by a power-plant is the rotational speed of this engine or motor. Table I gives the

TABLE I

Engine speed (R.P.M.).	Frequencies of predominant component (C.P.M.)	Order number (C.P.M./R.P.M.)
2070	5180	2·50
2320	5790	2·49
2540	6360	2·50
2630	6580 and 9200	2·50 and 3·50
2740	6860 and 9600	2·50 and 3·50
2880	10,100	3·51

frequencies of the predominant components in a set of waveforms recording the vibration of a certain aircraft engine on its test-bed; records were taken at the various speeds, given to the nearest

10 R.P.M. in the first column of the table. The frequencies were determined by comparing the waveform with a series of time indications marked on the same film by means of an alternating-current motor; this device marked a line on the film at intervals of 1/100 second, and by this means it was possible to determine the frequencies of the predominant components to the nearest 10 C.P.M. On dividing the frequencies by the corresponding engine speeds it was found that the large vibration motions were at frequencies equal to either $2\frac{1}{2}$ or $3\frac{1}{2}$ times the rotational speed of the engine; these results are given in the last column of the table. The slight divergences (in the case of the "order numbers" 2·49 and 3·51) can be accounted for as due to slight errors in the analysis. In this instance the information concerning the "order numbers" ($2\frac{1}{2}$ and $3\frac{1}{2}$) was all that was required; the actual frequencies were of no interest, and it would have been more convenient to have employed a device to put markings on the film corresponding to revolutions of the engine instead of hundredths of a second.

TABLE II

Engine speed (R.P.M.)	Order numbers.	Frequency of lower-order component (C.P.M.).
2070	$1\frac{1}{2}$ and 1·44	2980
2350	$1\frac{1}{2}$ and 1·28	3010
2630	$1\frac{1}{2}$ and 1·14	3000
2860	$1\frac{1}{2}$ and 1·05	3000
3050	$1\frac{1}{2}$ and 0·98	2980

Table II shows the order numbers of the predominant components of a set of waveforms taken in a test of another engine. In this test the time indication was in the form of an "engine tachometer trace," marking the completion of successive engine revolutions. The $1\frac{1}{2}$ order was quite pronounced and there was also a well-defined component whose order number varied from 1·44 at 2070 R.P.M. to 0·98 at 3050 R.P.M. On calculating the frequencies from the order numbers and engine speeds it was found (see last column of the table) that this second component was of a constant frequency of 3000 C.P.M., allowing for experimental error; in actual fact the component was traced to "hum" in the electrical amplifiers used in between the vibration pick-up and the recording oscillograph, 50 C.P.S. or 3000 C.P.M. being the frequency of the alternating-current mains supply of electricity.

These examples serve to show that it is sometimes convenient to employ one method of time indication, sometimes the other. Where possible it is useful to employ both methods simultaneously. Fig. 1 shows an actual recorded waveform where this has been done, the uppermost trace recording engine revolutions and the lowest trace recording time-intervals of 1/5 second. The detailed frequency analysis is given below, but it may be observed here that the inclusion of both revolution and time markings enables the speed of the engine to be checked very accurately. On the record illustrated, the distance between the engine markings 0 and 5 is 69·6 mm., while the distance between the markings E and F, representing

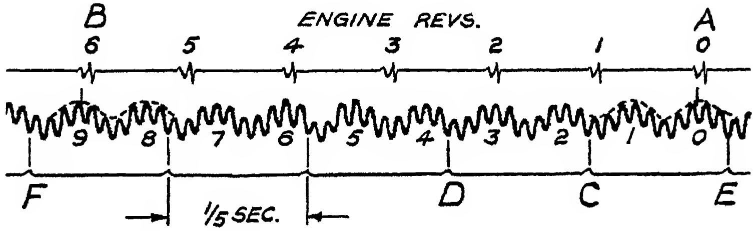


FIG. 1.—Record incorporating timing marks (EF) and engine revolution marks (AB).

one second, is 95 mm. The paper speed is therefore 95 mm./second. The engine speed is found thus :

$$\text{Time for 5 revolutions} = 69{\cdot}6/95 = 0{\cdot}733 \text{ second.}$$
$$\text{Time for 1 revolution} = 0{\cdot}733/5 = 0{\cdot}1466 \text{ second.}$$
$$\text{Engine speed} = 60/0{\cdot}1466 = 409 \text{ R.P.M.}$$

The same result is obtained by a slightly different method :

$$\text{One second is equivalent to } 95 \times 5/69{\cdot}6 = 6{\cdot}82 \text{ revolutions}$$
$$\text{Engine speed} = 6{\cdot}82 \times 60 = 409 \text{ R.P.M.}$$

It should be noted that the paper speed is sensibly constant over the length of the record, as shown by the even spacing of the timing marks ; it would be dangerous, however, to assume that the speed remains constant from record to record, particularly as in some instruments the speed depends upon the amount of paper left on the spool. The speed should always be checked afresh in every record, unless special provisions have been made to ensure that it remains constant.

This process of matching one trace against a standard time-base trace is fundamental to the determination of frequencies. Where definite cycles of the components concerned can be seen in the waveform, the length of record occupied by some convenient number

of cycles of the component is measured, and so is the length of record covered in a certain time or a certain number of engine revolutions (or whatever may be the time-base). The frequency is determined in a manner similar to that employed in the engine speed check described above. Before proceeding to consider examples of this process, it must be pointed out that accuracy in frequency determination depends upon two factors: the constancy of the paper speed of the record, and the possession of sufficient length of the record. The first of these requirements can in certain conditions, shortly to be considered, be dispensed with, but the second is a general requirement. Fig. 2 shows parts of the same waveform with a time indication included. If the short record (Fig. 2*a*) is all that is available for analysis, it appears that there are four

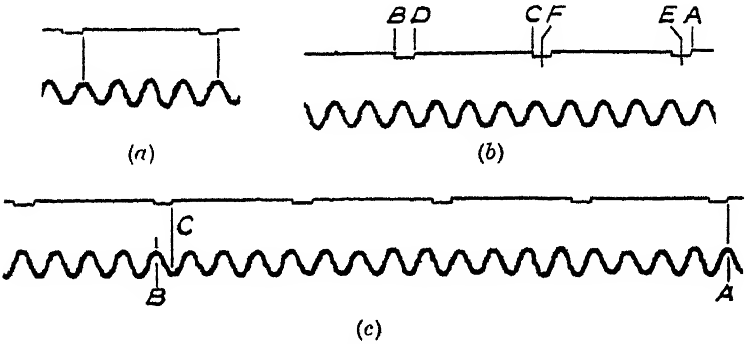


FIG. 2.—Illustrating the importance of utilising a fairly long record in determining frequencies. From (*a*) it appears that four cycles of the wave occur in one timing interval, but from (*c*) it is clear that $16\frac{1}{2}$ cycles occur in four timing intervals.

cycles of the wave to one timing interval. If the longer record (Fig. 2*c*) is available, however, it can be seen that there are in fact $16\frac{1}{2}$ cycles to four timing intervals. This is clearly a more accurate assessment, and taking it to be correct the relative accuracy of the first assessment is $4 \times 4/16\frac{1}{2} = 0{\cdot}97$, showing an error of 3 per cent. The timing interval in this case is 0·82 seconds, obtained from a governed clockwork motor. The true frequency is determined as follows:

$$16\tfrac{1}{2} \text{ cycles occur in } 4 \times 0{\cdot}82 = 3{\cdot}28 \text{ seconds.}$$
$$1 \text{ cycle occurs in } 3{\cdot}28/16\tfrac{1}{2} = 0{\cdot}199 \text{ seconds.}$$
$$\text{Frequency} = 1/0{\cdot}199 = 5{\cdot}03 \text{ C.P.S.}$$
$$\text{or } 60 \times 5{\cdot}03 = 302 \text{ C.P.M.}$$

The frequency calculated from the inaccurate estimate based on the short record (*a*) would be $60 \times 4/0{\cdot}82 = 293$ C.P.M.

Another point concerning the use of time markings requires mention: with a marking of the type shown in Fig. 2, it is essential always to use the same part of the mark. In (*a*) and (*c*) the right-hand part of the mark has been used. It is obviously inaccurate to measure from the right-hand end of one mark to the left-hand end of another, as from A to B in Fig. 1*b*, or *vice versa* as from C to D: in the first case too high, and in the second case too low, a value will be found for the frequency. The estimation of the midpoint of the marking, as at E and F, can also lead to errors, although such errors may only be slight. It is much the best plan to use some definite part of the mark, such as the right-hand corner.

The waveform in Fig. 3 exhibits a non-uniform time-base, shown up by the uneven spacing of the markings on the timing trace at the bottom of the record. Nevertheless it is possible to determine the fact that 9½ cycles of the vibration occur in four

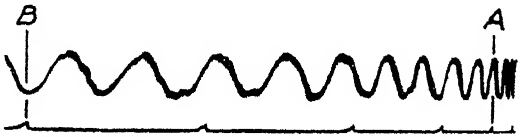


FIG. 3.—Accelerated record, showing non-uniform time-base.

timing intervals (the part AB of the record). These timing intervals are 1/10 second, and the frequency is determined as

$$60 \times 9{\cdot}5 \times 10/4 = 1425 \text{ C.P.M.}$$

The record shows vibrations in an airframe when the engine is running at 3000 R.P.M. (nominal value), and there are two possible sources of the vibration—the engine and the propeller. The gear-ratio being 0·477 : 1, the nominal propeller speed is 1440 R.P.M. It so happens that in this case the vibration frequency must be equal either to an integral multiple of the propeller rotational speed or to an integral multiple of half the engine rotational speed. The frequency found from the record (1425 C.P.M.) shows that if the vibration is at 1 × propeller R.P.M. the true engine speed is 2990 R.P.M., whereas if the vibration is at ½ × engine R.P.M. the true engine speed would be 2850 R.P.M. Since it is known that the tachometer by means of which the nominal engine speed was obtained is accurate to about ± 1 per cent. at 3000 R.P.M., the true speed must lie between 2970 and 3030 R.P.M. The analysis indicates that the true speed is 2990 R.P.M. and the vibration is at 1 × propeller R.P.M.

From the waveform in Fig. 4 it is required to determine the order numbers of the components referred to engine R.P.M., the

timing trace marking intervals of one second and the engine speed being 200 R.P.M. (nominal). The order numbers must belong to the series $\frac{1}{2}$, 1, $1\frac{1}{2}$, 2, $2\frac{1}{2}$, etc., × engine R.P.M. The distance AB, representing two seconds, includes $16\frac{3}{4}$ cycles of the high-frequency component. The frequency can therefore be obtained directly as $60 \times 16{\cdot}75/2 = 503$ C.P.M. Alternatively, by measurement along the record, AB = 63 mm. (equivalent to 2 seconds) and CD = 64 mm. (equivalent to 17 cycles). The paper speed is therefore 31·5 mm./second, and the frequency is $60 \times 17 \times 31{\cdot}5/64 = 502$ C.P.M. Since it is known that the frequency must be $n/2 \times$ engine R.P.M., where n is an integer, it is seen that in this case $n = 5$ and the true engine speed is $502 \times 2/5 = 201$ R.P.M.

In this wave there are two components, the ripple at 502 C.P.M. being superimposed upon a lower frequency variation ; the general form of the record is very similar to that illustrated in Chapter I, Figs. 18-20, and recalling the results obtained in Section 10 it is

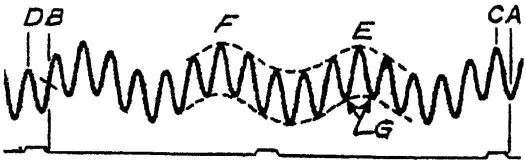


FIG. 4.—Two-component wave (frequency analysis): the fact that the symmetrical disposition of the troughs " G " is repeated throughout the wave is an important clue.

evident that the lower frequency component is represented by the " envelope " drawn in a broken line in the centre part of the record. The distance EF, containing one cycle of the lower frequency component, contains exactly five cycles of the 502 C.P.M. component. Since the low frequency must be $n/2 \times$ engine R.P.M., where n is an integer, it is apparent that in this case $n = 1$.

The following information has been extracted from the record :

True engine speed	= 201 R.P.M. to nearest unit.
Low frequency	= 100 C.P.M. to nearest unit
	= $\frac{1}{2}$ × engine R.P.M.
High frequency	= 502 C.P.M. to nearest unit
	= $2\frac{1}{2}$ × engine R.P.M.

The amplitudes of the two components can be found quite simply and the process is described in Section 3 below.

In this particular case the data on which the conclusions are based consist of the record, together with the fact that the frequencies

must belong to the series $\frac{1}{2}$, 1, $1\frac{1}{2}$, etc., engine R.P.M. If this additional information is not available and the length AD of the record is the only basis for argument, the frequencies of both components can still be obtained with the same accuracy. The only assumption requiring justification is that the lower frequency is exactly 1/5 of the higher frequency. A glance at the record shows that the frequency ratio is certainly 1 : 5 approximately; closer study reveals the fact that the two troughs marked G are symmetrically disposed with regard to the crest E, and the same is true at all parts where the low-frequency component has crests. There is no tendency for the high-frequency ripple to "creep" along the low-frequency component, and this shows that there is a simple integral relationship between the two frequencies. Naturally, the possession of a longer record would enable this fact to be demonstrated with greater rigour.

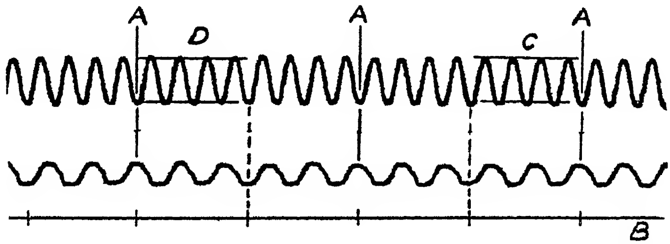


FIG. 5.—Use of a sine-wave of known frequency as a time-base reference. The lower trace is known to have a frequency $2\frac{1}{2}$ times engine R.P.M., and the scale B, showing revolutions, can be derived.

In Fig. 5 a double waveform is shown. The upper trace records a vibration of (as yet) undetermined frequency, and the lower trace is known to represent a vibration of the $2\frac{1}{2} \times$ engine order, i.e. its frequency is $2\frac{1}{2} \times$ engine R.P.M. The two traces were recorded simultaneously on a double-indicating instrument; on comparing them it is seen that the intervals between the successive points A, containing 5 cycles of the $2\frac{1}{2} \times$ order and therefore representing two engine revolutions, contain 8 cycles of the top trace. This trace therefore represents a vibration at $4 \times$ order, i.e. at a frequency $4 \times$ engine R.P.M. Alternatively, the artificial "engine revolution scale" B can be constructed, since it is known that $2\frac{1}{2}$ cycles of the lower trace occur in one engine revolution. It is then seen directly that the frequency of the upper trace is $4 \times$ engine R.P.M.

The examples described show how the frequency of a wave can be determined either directly or as a multiple of some standard frequency, by means of (i) a time-recording trace, (ii) a trace

recording the standard frequency, or (iii) a trace containing a component at a known frequency. By obvious extensions of the method it is possible to determine the frequencies of all the components in a waveform provided that the frequency of one component has been determined and the wavelengths of all the components can be found.

It is customary, and advisable where possible, to utilise a separate trace recorded simultaneously with the main trace, for the purpose of time indication. In some recording equipment, where the pen or light-spot which actually makes the record is controlled electrically, a timing mark is superimposed on the main trace. Examples of this type of time indication are given in Fig. 6. The "zig-zag" at A in Fig. 6*a* indicates the start of the firing cycle in an internal combustion engine; the terminals of the galvanometer used for recording the main trace are connected also to a coil surrounding the ignition lead to the sparking-plug in one cylinder, and this

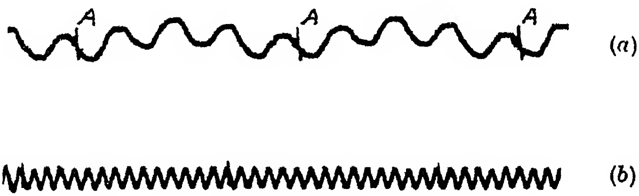


FIG. 6.—Superimposed timing marks: (*a*) on a fairly "open" wave; (*b*) on a high-frequency ripple.

auxiliary circuit receives a "kick" in current by induction every time this cylinder fires, i.e. once every two engine revolutions. This method of recording time indications is not to be recommended: apart from the practical difficulties (which include the avoidance of "pick-up" from the ignition on other cylinders) the mark itself is not easy to see. In the record shown in Fig. 6*b* there are three such marks, and the reader will observe the difficulty of "spotting" them. With very high-frequency components present in the waveform the difficulty increases, and although the objection is not insuperable when only a small number of records have to be analysed, the eye-strain imposed on the analyst by the employment of this type of time marking can become a serious matter when any large number of records is handled.

With the double-beam cathode-ray tube, and certain other recording instruments, it is convenient to use a 50 C.P.S. trace for frequency calibrations. Such a trace is easily derived from the A.C. electric mains supply.

A general formula can be given for the determination of frequencies. If l mm. is the distance between two timing marks corresponding to a time-interval of t seconds, the time-base factor or paper speed K_t is given by

$$K_t = l/t \text{ mm./second,} \qquad (2.1)$$

and the time t seconds corresponding to any distance l mm. of the trace is given by

$$t = l/K_t \text{ seconds.} \qquad (2.2)$$

If λ mm. is the wavelength of the component whose frequency F is being determined, then from (2.2) the period T seconds of the component is given by

$$T = \lambda/K_t \text{ seconds.}$$

Now $F = 1/T$ C.P.S., hence

$$F = K_t/\lambda \text{ C.P.S.,} \qquad (2.3)$$

or, substituting from (2.1),

$$F = l/\lambda t \text{ C.P.S.} = 60l/\lambda t \text{ C.P.M.} \qquad (2.4)$$

In this formula, as stated in the derivation, F = frequency of any component, λ = wavelength of that component in mm., and l = distance in mm. corresponding to t seconds.

If a length L mm. of the record contains p cycles of the component, the wavelength λ is given by

$$\lambda = L/p \text{ mm.} \qquad (2.5)$$

The distances L, l and λ may of course be measured in inches, or indeed in any arbitrary units, without affecting the result; but they must be measured in the same units.

If engine revolutions or some other frequency standard is employed, the calculation may be expressed thus: let l' mm. be the length of record between two marks whose separation represents m cycles of the frequency standard whose frequency is N C.P.S. The time-base factor is then

$$K_t = l'N/m \text{ mm./second,}$$

and from (2.3),
$$F = \frac{l'N}{\lambda m} \text{ C.P.S.} \qquad (2.6)$$

Moreover, if n is the order number of the required frequency referred to the frequency standard,

$$n = \frac{F}{N} = \frac{l'}{\lambda m}. \qquad (2.7)$$

In the case of engine-excited vibrations, N would represent the engine speed in revolutions per second; since it is more usual to give such speeds in R.P.M., the formula (2.6) can be used in these cases with N in R.P.M. so long as it is remembered that F will then be given in C.P.M. instead of C.P.S.

The principles outlined above may be applied to the determination of the fundamental frequency of a complex wave.

The application of these formulæ can be illustrated by the determination of the frequencies in the waveforms in Figs. 1-5, and some of these analyses are given in detail below.

Fig. 1—*high-frequency component.*

(*a*) Time-base.

$$\text{Distance EF} = l = 95 \text{ mm.}$$
$$\text{Time represented by EF} = t = 1 \text{ second.}$$
$$\text{Wavelength} = \lambda = \text{EF}/50 = 95/50 \text{ mm.}$$
$$\text{From (2.4) frequency} = \frac{60 \times 95 \times 50}{95 \times 1} = 3000 \text{ C.P.M.}$$

(*b*) Order number.

$$\text{Distance AB} = l' = 83{\cdot}9 \text{ mm.}$$
$$\text{AB contains } m = 6 \text{ engine revolutions.}$$
$$\text{From (2.7) order number } n = \frac{83{\cdot}9 \times 50}{95 \times 6} = 7{\cdot}35 \text{ cycles/revolution.}$$

(*c*) $\text{Engine speed N} = \dfrac{\text{frequency}}{\text{order number}} = 3000/7{\cdot}35 = 409 \text{ R.P.M.}$

Fig. 1—*low-frequency component.*

The envelope drawn in broken lines at each end of the record displays the low-frequency component, nine cycles of which occur between the two short vertical lines above the figures 0 and 9. The distance is 84·2 mm., which is sufficiently close to the value 83·9 mm. for l' to show that the nine cycles occur in six revolutions—i.e. the order number is $9/6 = 1\frac{1}{2}$.

Fig. 2*c*.

Distance AB = L = 78·5 mm., containing $p = 17$ cycles of wave.
Wavelength $\lambda = \text{L}/p = 78{\cdot}5/17$ mm.

Distance AC = l = 76·1 mm., representing $4 \times 0{\cdot}82$ seconds,
i.e. $t = 3{\cdot}28$ seconds.

$$\text{From (2.4), F} = \frac{60 \times 76{\cdot}1 \times 17}{78{\cdot}5 \times 3{\cdot}28} = 302 \text{ C.P.M.}$$

Fig. 4—high-frequency component.

Distance AB $= l = 63$ mm., representing $t = 2$ seconds.
Distance CD $= L = 64$ mm., containing $p = 17$ cycles.
Wavelength $\lambda = L/p = 64/17$ mm.

$$\text{From (2.4), } F = \frac{60 \times 63 \times 17}{64 \times 2} = 502 \text{ C.P.M.}$$

The formulæ are reprinted below for ease of reference :

Let L mm. = length of record containing p cycles of component,
l mm. = length of record representing t seconds,
l' mm. = length of record containing m cycles of frequency N C.P.S.,
λ mm. = wavelength of component,
K_t mm/second = time-base constant or paper speed.

Then $\lambda = L/p$, $K_t = l/t = l'N/m$,
and $F \text{ (C.P.S.)} = K_t/\lambda = l/\lambda t = lp/Lt = l'N/\lambda m = l'pN/Lm$ (2.8)

3. Determination of amplitudes in simple cases.

The measurement of the amplitude of a sine-wave requires little explanation. The amplitude is simply half the total variation in height of the wave, as shown in Chapter I, Fig. 2, p. 9.

When great accuracy is required, the thickness of the trace must be taken into consideration. The usual method of determining amplitudes, when there are many traces to be analysed and accuracy is required, is to draw with a fine pencil two horizontal lines passing respectively through crests and troughs. These lines may be drawn through the estimated centre of the trace thickness, as at D in Fig. 5, or both to one side of the trace ; at C in this diagram the lines have been inscribed so as to touch the top edge of the trace, both at the crests and at the troughs. The vertical distance between the two lines gives the double amplitude of the wave directly ; in the diagram, this distance is 5·9 mm., so that the amplitude is 2·95 mm. It is customary and advisable to take several measurements at various parts of the record and average them.

A little instrument, known as the "Scaleometer," is very useful for obtaining accurate amplitude measurements. It consists of a simple magnifying system through which the trace and a finely-divided scale are viewed, the scale being engraved on a glass disc in such a way that it can be brought down on to the trace. Special scales are available to measure in small subdivisions of either inches or millimetres.

True amplitudes obtained in this manner usually require to be multiplied by some constant factor in order to take into account the calibration constant or magnification of the recording system (see Chapter II, p. 49). This being the case, it is frequently possible to make use of double amplitude readings, which are easier to measure than true amplitudes (in so far as no calculation is required —even division by 2 is time consuming and can lead to errors), the conversion factor being altered accordingly. For example, suppose a trace recording an alternating-current voltage has a double amplitude of 12·2 mm., the sensitivity of the recording system being such that 3·2 mm. excursion on the trace corresponds to a voltage variation of 1 mv. The total voltage variation is then $12{\cdot}2/3{\cdot}2 = 3{\cdot}82$ mv., representing a variation between $+ 1{\cdot}91$ and $- 1{\cdot}91$ mv. This is known to electrical engineers as a wave of 1·91 mv. peak; and this peak voltage figure could have been obtained directly from the record, using the true amplitude 6·1 mm. instead of the double amplitude. Alternatively, the calibration constant can be taken as $2 \times 3{\cdot}2 = 6{\cdot}4$, signifying "6·4 mm. double amplitude for 1 mv. peak," and the peak voltage can then be calculated from the double amplitude as $12{\cdot}2/6{\cdot}4 = 1{\cdot}91$ mv. peak, as before.

The same remarks apply to the determination of stress amplitudes from strain-gauge records. It is usual to express vibratory stress levels in such a form as "$\pm$ 8000 lb./in.^{2}," signifying an overall variation of 16,000 lb./in.^{2}, even when this variation is not symmetrically disposed with respect to the mean or centre position *of the recording pen or light-spot;* strain-gauges are usually electrical in nature, and the steady stress or mean stress value is not recorded, since it does not "register" in the intermediate amplifiers and other electrical apparatus.

Throughout the remainder of this book the two types of amplitude measurement will be distinguished thus:

True amplitudes will be prefaced by the sign "$\pm$."

Double amplitudes will be followed by the letters "(d.a.)."

Thus the amplitude of the trace in Fig. 5 is $\pm$ 2·95 mm. or 5·9 mm. (d.a.).

In many cases in engineering practice it is required to measure the overall excursion of complex waves. This occurs, for example, in the determination of vibratory stresses by means of strain-gauges. The waveform in Fig. 7 is an example of the type of record frequently encountered. The length AB contains one cycle of the stress variation. The line AB, joining corresponding peaks in successive cycles, is parallel to the mean line of the trace, so that measurements taken perpendicular to AB will be true indications

of amplitudes. At C the trace reaches the lowest position below AB, and the distance from C to the line AB (conveniently denoted by C—AB), is the true double amplitude of the wave; that is, the stress certainly varies through the range represented by this distance once per cycle. It has been the practice of some analysts, engaged in the routine analysis of many hundreds of such records, to measure a distance such as E—AB, or the distance between the lines C and D, on the assumption that the slightly greater distance C—AB is the result of an extraneous low-frequency surge. This is a case which admirably illustrates the general rule that the precise technique adopted in any analysis will depend upon the interpretation which is to be put upon the results and upon a knowledge of the physical conditions governing both the system under observation and the recording system. The wave shown in the diagram is perfectly periodic, the wavelength being 150 mm. in the original, so that a low trough is found 150 mm. to either side of C, and so throughout the record. This being the case, the so-called "extraneous low-frequency surge" is an inherent part of the total variation and should not be omitted from the total stress estimate. In other words, the total periodic stress variation is over the range represented by the distance C—AB (8·7 mm. in the original); the measurement E—AB or C—D is only 8·1 mm. in the original, giving an error of nearly 7 per cent. The trace records vibratory stresses in an aircraft propeller, the engine having a reduction gear-ratio of 0·4 : 1. The full cycle of the variation extends over ten engine revolutions (see Chapter II, p. 47) since disturbances arise at frequencies which are multiples either of the propeller speed or of half the engine speed. The predominant components are seen to have 20 and 4 cycles respectively in these ten engine revolutions, and represent vibrations at 2 × engine R.P.M. (2E) and 1 × propeller R.P.M. (1P) respectively; there is also present a number of minor components which have the effect of extending the wavelength of the total variation to the maximum possible value, equivalent to ten engine revolutions

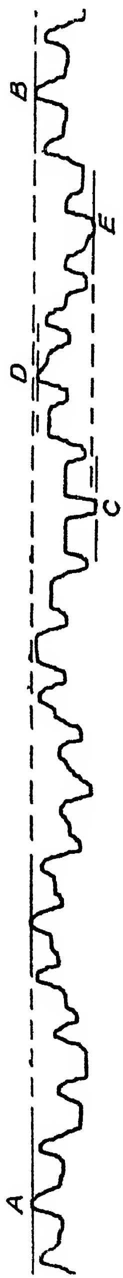


FIG. 7.—Complex strain-gauge record: AB is a cycle.

or four propeller revolutions. It is admittedly true that the full variation (8·7 mm.) occurs only once every cycle, whereas the variation E—AB (8·1 mm.) occurs four times in the cycle; nevertheless, 8·7 mm. is the correct value for the overall variation so far as the waveform analyst is concerned; the interpretation of the results is a matter for the vibration engineer to determine.

The procedure adopted here for the determination of the correct direction in which to measure amplitudes is one of universal application: find corresponding points in successive cycles of the wave, and join them; the line so obtained is parallel to the mean line of the trace, and amplitude measurements taken perpendicular to it will be correct. The point is of some importance in the measurement of the overall variation in complex waves, whose wavelengths may be as great as twenty or thirty centimetres in some cases.

Some recording instruments are so arranged as to indicate the mean line down the centre of the trace, as for example in Figs. 22 and 23, Chapter I. In these diagrams it can be seen that although the mean line represents the average value of the variation over a cycle, it often happens that the maximum height of the wave above this line is greater or less than the maximum depth of the wave below it. In Fig. 22*e*, for example, the maximum height of the wave above the mean line is only about half the maximum depth below it. In the original the two distances were respectively 0·16 and 0·30 in.; nevertheless, since the constant term in the Fourier series (see Chapter II, p. 49) has in this case been removed by the electrical apparatus intervening between the source of the variation and the recording equipment, the amplitude would be stated as $\pm$ 0·23 in.

4. Phase-angle determination in sine-waves.

The determination of the phase-angle of a sine wave is a simple matter, which is effected by finding the precise location of the crests and troughs, or alternatively of the zero-points, where the trace cuts the mean or centre-line. Suppose that it is required to express the variation, represented by the sine-wave in Fig. 8*a*, in the form $y = a \,.\, \sin(\omega t + \phi)$, the line A being the time-base datum $t = 0$, and the frequency being 3000 C.P.M. From formula (3.1) on page 12, $\omega = 3000/9{\cdot}55$, and the constant a is simply the amplitude of the sine-wave. The phase-angle ϕ is found as follows: the lines B and C are drawn, passing respectively through crests and troughs (Fig. 8*b*). The line D, which is the mean line of the trace, is midway between the lines B and C, and so is easily constructed. Let the mean line cut the time-datum A at E, and let

F be the intersection nearest to A and to the left of A (i.e. *before* A in the sense of time indicated by the arrow) of the mean line D and the trace where the slope is positive (i.e. where y is increasing with time). FG is one cycle of the trace. Now, the distance FG represents an angle 2π radians or 360°, and on the same scale the distance FE represents ϕ; for at F, $y = 0$ and at E, $y = a \,.\, \sin\,(\mathrm{FE}/\mathrm{FG})2\pi$.

In the particular case shown, FG = 1·2 in. in the original and FE = 0·48 in.; thus $\phi = (0{\cdot}48/1{\cdot}2)2\pi = 0{\cdot}8\pi$ radians or 144°. If the trace had not extended to the left of the line A, the phase-angle

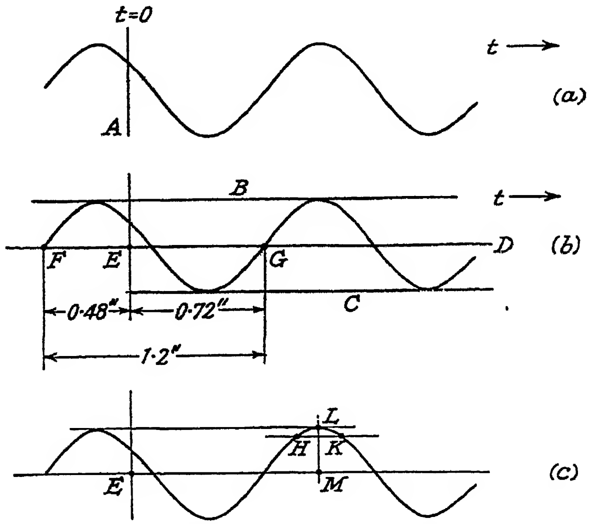


FIG. 8.—Technique of phase-angle determination.

could have been found as $\phi = -(\mathrm{EG}/\lambda)2\pi$, where λ is the wave-length of the trace, for

$$(\mathrm{FE} + \mathrm{EG})/\lambda = 1 \text{ and } \sin\,(\omega t + \phi) = \sin\,[\omega t - (2\pi - \phi)].$$

Thus in Fig. 8*b*, EG is found to be 0·72 in., and

$$\phi = -(0{\cdot}72/1{\cdot}2)2\pi = -1{\cdot}2\pi \text{ radians}$$

or − 216°; and since

$$216° + 144° = 360°,\ \sin\,(\omega t - 216°) = \sin\,(\omega t + 144°).$$

In order to express the variation in the form

$$y = a \,.\, \cos\,(\omega t - \psi),$$

use may be made of the fact that if

$$\sin\,(\theta + \phi) = \cos\,(\theta - \psi),$$

then
$$\psi = \phi - \pi/2;$$

for $\cos(\theta + \phi - \pi/2) = \sin(\theta + \phi)$, as shown in (2·5) in Chapter I, page 8. Thus in the case illustrated, $\psi = 144° - 90° = 54°$.

The method just described for the determination of phase-angles is the most accurate method. For rough approximations the location of the crests and troughs can be estimated by eye. Another accurate method consists of drawing a line parallel to the crest line B and a short distance below it, cutting the trace either side of a crest at H and K (Fig. 8c). A short line perpendicular to and bisecting the intercept HK cuts the trace at the crest L and the mean line at M; then the fractional part of the ratio EM/λ, expressed as a proper fraction, when multiplied by 2π gives the value of $(-\psi)$ in the form $a \,.\, \cos(\omega t - \psi)$. This same method of utilising the fractional part of the ratio when the length EM exceeds λ can, of course, be used in the first procedure described above.

When the mean line has been constructed, the sine-wave trace cuts off a series of equal segments on it. If the segments are seen to be unequal, being alternately longer and shorter than half a wavelength, some distortion is present in the waveform. It does not, of course, follow that the presence of equal segments indicates that the trace is a pure sine-wave; but if unequal segments are found, and the mean line has been accurately constructed, it can at once be deduced that the wave is not a pure sine-wave and requires analysis into its components before any phase-angles can be determined.

In the following chapter, where the analysis of complex waves having two or more sine-wave components is described, the methods of finding phase-angles and the location of crests and troughs described above will be understood to be employed, with certain modifications according to the particular circumstances of their employment.

CHAPTER IV

THE ENVELOPE METHOD

1. Introductory.

The envelope method of inspection analysis, described in this chapter, has been evolved by vibration engineers for the routine analysis of large numbers of recorded waveforms. Large-scale vibration tests, particularly those in which stress values are recorded by means of strain-gauges and auxiliary equipment, usually comprise several " runs " over a wide range of speeds of the power-plant under test, records being taken from numerous vibration pick-ups. As a result, each test may provide some hundreds or even thousands of traces, and from each trace the following information is required : (i) the overall amplitude of the variation ; (ii) the frequencies of the principal components, and (iii) their amplitudes. Phase-angles are not usually required to be determined in every trace of a test, but they may be required in certain typical traces.

The order in which these requirements are listed indicates the descending order of importance for most practical requirements. Consider, for example, stress determinations by means of strain-gauges. Of prime importance is the overall variation in the stress, which, together with the level of the direct stress, determines the fatigue characteristics of the structure under test ; the process for measuring this overall amplitude has been described in Chapter III. Secondly, the frequencies of the principal components must be found before means of reducing the vibration stresses can be sought, since the source of each component must be located. Thirdly, the approximate amplitudes of each component must be determined, so that if means of reducing one component are found the effects on the overall amplitudes of the various records can be estimated. In the accumulation of a " background of experience " (which is the research worker's most useful acquisition) the phase-relations between different records is often of great importance ; also, in the strain-gauge investigation of stresses (notably in aircraft propellers) these phase-relations may enable otherwise doubtful frequencies to be found more accurately, particularly where for some reason the record is too short.

The envelope method of waveform analysis enables these different quantities to be determined with sufficient accuracy for

all practical purposes. It is true that there are certain special types of wave, containing only three components, which cannot be analysed by the method, but examples of these types occur but seldom in practice, and even when they do occur it is usually possible to determine the frequencies of the components by a consideration of records taken at speeds slightly different from that of the record in question, and when several traces are recorded simultaneously there is usually one trace wherein one of the components is not present. The vast majority of waveforms encountered in engineering vibration study fall into one of a small number of general types, to which a systematic method of analysis is applicable.

The first chapter of this book is concerned with the properties of sine-waves in combination; beats, and the effects of adding low-frequency or high-frequency components to existing combinations are discussed therein. The general properties of these combinations form the basis of the envelope method of analysis. The technique of analysis could be justified mathematically, but the absence of any necessity for such justification is sufficient reason for not describing here the complicated mathematical work. The technique is based upon perfectly general properties of sine-waves in combination, which have been illustrated in Chapter I by several examples; and the method has been fully tested in a leading industrial establishment, where many thousands of recorded waveforms have been analysed by its aid.

Practice in applying the method leads to great facility; experienced analysts are able to estimate "at a glance" the principal components of a waveform analysable by the method. It must be emphasised, however, that there is no royal road to expert analysis: facility is only to be obtained by practice, and the best place to start is the synthesis of complex waves from their components. There can be no doubt, indeed, that time so spent is by no means wasted, since the analyst is afterwards in a position to perform the required estimations with a sufficient degree of accuracy and in the minimum of time. The possibility of anomalies in recorded waveforms should always be borne in mind; the more waveforms an individual analyses, the greater will be his experience and understanding of the idiosyncrasies of different combinations.

In the present chapter the analysis of waveforms with only two components is first described, the technique being developed from properties of two-component waves recalled from Chapter I. Various types of three-component waveforms are next discussed, the technique being a logical development of that used for two-component waves. From a study of the methods of treatment of

these various types of wave, a systematic scheme of analysis can be derived, and this is given in the text in the form of two tables. The application of these general tables is illustrated by means of actual recorded examples previously analysed in the chapter; the last example is of a waveform with four predominant components, and shows the limit of application of the method. The chapter concludes with some notes on practical details.

2. Determination of cycle, and apparent highest frequency.

The first task to be performed in the analysis of any waveform is the determination of the cycle. Since the frequencies of all the components of a complex wave are multiples of the frequency of the fundamental cycle, the determination of the cycle renders the frequency analysis much easier. In many cases the cycle is very clearly evident, and the majority of the waveforms illustrated in this chapter are of this type; in certain cases, however, care must be taken not to regard too short a distance as being the wavelength. The need for the exercise of care in this connection is exemplified in the trace of Fig. 1. At first sight the cycle appears to be from A to B, or from B to A′, so that the length AA′ contains two cycles. More careful inspection reveals, however, that the form of the waist of the beat at B is not quite the same as at A or A′; in fact, the trace is approximately alternant, with half-cycles AB and BA′. Only a portion of the record is shown in the diagram, but in the original trace the pattern AA′ was repeated accurately throughout, waists of type A alternating with those of type B.

In order to determine accurately the length of the cycle, use may very effectively be made of "characteristic patterns." Thus, when the form of the trace in the neighbourhood of A in the diagram has been observed, it will at once be evident that AB is only part of the cycle, as the trace at B does not conform to this characteristic pattern.

In many waveforms recording mechanical vibrations there are irregular fluctuations from cycle to cycle. [Strictly speaking, the term "cycle" is inapplicable in such a case, but where the fluctuation is small, as it usually is, it is quite justifiable to neglect the effects of this irregularity.] The variation from cycle to cycle is normally in the size of the various peaks, the location of the peaks being for all practical purposes the same in each cycle. Some definite pattern, regularly repeated, can usually be observed. If the type B waist in Fig. 1 had occurred only once in a dozen or so beats, all the other waists being of type A, and if each beat contained the same number of peaks, the occurrence of the type B waist would

be considered as due to irregular fluctuation; in the actual record from which the diagram is reproduced, however, the types A and B alternate and there is no doubt that the cycle is AA′.

In the following sections the term "harmonic" refers to this cycle of the waveform, and not to any other frequency standard. Thus if the cycle contains thirteen cycles of a particular component, that component will be termed the 13th harmonic. In some cases, the harmonic number may require to be converted to another number showing the relation between the frequency of the component and some standard frequency: thus with waveforms recording vibrations in engines, the frequencies are usually expressed as multiples of the crankshaft rotational speed, and this frequency ratio is frequently termed the harmonic number; nevertheless, it seems to be more convenient to reserve this term to show the relation between the frequency of the component and the fundamental frequency of the variation. If the waveform contained only the 2E and 6E engine orders, i.e. components at two and six times

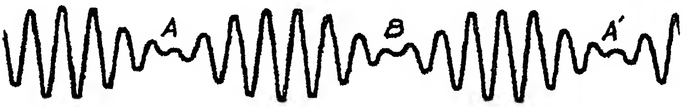


FIG. 1.—Use of characteristic patterns in determining cycles. The cycle is AA′, not AB, as the form of the waist at B is not the same as at A, A′. For analysis see Table IV, page 114.

engine R.P.M. respectively, the cycle of the variation extends over only half an engine revolution and the components are the first (or fundamental) and third harmonics of this cycle; similarly, the $\frac{1}{2}$E order is the fundamental component in the firing cycle of a 4-stroke internal combustion engine. It is convenient and desirable to retain the term "harmonic" with reference to the cycle of the waveform, and to use the term "order" or "order number" with reference to auxiliary frequency standards.

Once the cycle has been identified, the frequency of the fundamental component can be determined (whether or not it is present in the waveform) by the methods described in Chapter III, and the frequencies of the other components are found by multiplying their harmonic numbers by the fundamental frequency.

The determination of the frequencies of the various components is therefore straightforward, once it is established how many cycles of each component occur in the fundamental cycle. The methods by which these numbers are established are described in the following sections, but one point concerning the apparent highest frequency

requires immediate mention. As shown in Chapter I, page 20, there cannot be more peaks in a cycle of a complex waveform than there are in a corresponding length of the component of highest frequency. If, for example, there are eighteen crests in a cycle, then there is certainly present a component whose harmonic number is not less than eighteen. There may be present components of higher frequency, in the case of beating waveforms; but confining attention here to the highest frequency apparent in the trace, some of the peaks even of this component may be masked. At A in

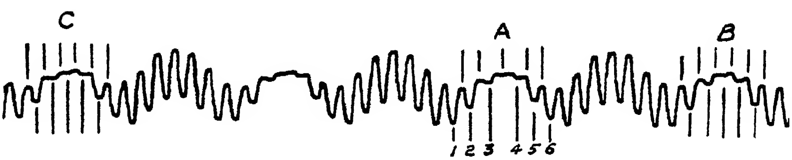


FIG. 2.—Masking of peaks of high-frequency component.

Fig. 2, for example, the crests and troughs are apparently as indicated by the vertical lines; but this entails a much wider separation between successive troughs "3" and "4" than at either side. The true state of affairs is as indicated at B and C, where in fact the peak separation at the waist of the beat is slightly less than elsewhere. As is explained below (p. 94) the marking of peaks in beating waveforms does not affect the results: in this particular case, the correct spacing of the peaks as at B indicates the presence of a 13th harmonic beating with a 12th, while the incorrect spacing

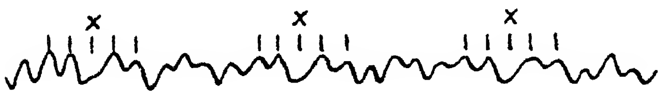


FIG. 3.—Masking of peaks. Peaks at the points marked X are masked in the record.

as at A indicates a 12th harmonic beating with a 13th; and since the minimum width of the envelope strip (at the waist of the beat) is zero, the two components have equal amplitude. It is therefore immaterial which assumption is made concerning the peak separations. On the other hand, with a single high-frequency component it is important to locate all the peaks. In Fig. 3 the crests marked X are masked in the trace, but they must be taken into account in determining the frequency of the high-frequency component. Bearing in mind the slight variations in peak spacing in beating waveforms, the masking of peaks is shown up by large variations

in the spacing. The practical rule is to insert peaks where necessary to preserve a more-or-less even distribution.

3. Two components of high frequency-ratio.

A complex wave consisting of two components whose frequency-ratio is high is characterised by the following properties (see Chapter I, p. 30):

(i) The envelopes, passing respectively through all the crests and all the troughs which can be seen in the trace, or touching the trace near these peaks, are similar; and each is a sine-wave.

(ii) The envelope strip, or area between the top and bottom envelopes, is constant in width, i.e. the vertical separation of the two envelopes does not vary over the cycle.

(iii) The high frequency "ripple" fills in the envelope strip, and has the appearance of a sine-wave distorted so as to lie between the two envelopes.

[These characteristics are illustrated in the waveforms of Figs. 18-20, Chapter I, pp. 29, 30].

The envelopes represent the low-frequency component, and the constant vertical separation of the envelopes represents the double amplitude of the high-frequency component. The construction of the envelopes therefore provides complete information concerning the low-frequency component, whose amplitude, frequency and phase can be determined by the methods described in Chapter III; the amplitude of the high-frequency component is also known. The determination of the frequency and phase of this wave depends upon the precise location of the peaks, which will now be considered.

Fig. 4*a* shows a wave whose two components have a high frequency-ratio. The envelopes, drawn in broken lines, are reproduced below at (*b*), and represent in all respects the low-frequency component. In the original wave the wavelength was 20 mm. and the double amplitude 4·2 mm. The double amplitude of the high-frequency component, represented by the vertical separation of the lines drawn at A, was 4·8 mm. BC comprises one cycle of the complete variation, and the frequencies of the two components can therefore be determined quite simply: they are the 3rd and 16th harmonics. All the peaks of the 16th harmonic appear in the trace. Calculation of the frequency of this harmonic by measuring the horizontal interval occupied by a number of its cycles is not to be recommended, since peaks are displaced slightly

in a horizontal direction by the superposition of the low-frequency component. This fact is illustrated in Fig. 5*a*; at (*b*) in Fig. 5 the crest of the high-frequency component coincides with a trough of the low-frequency component, and there is no horizontal displacement. Frequency and phase determinations are therefore best made by means of crests and troughs which coincide with peaks in the low-frequency envelope: the phase of the 16th harmonic in Fig. 4 is given precisely by the fact that a crest occurs at B, together with a crest of the 3rd harmonic. These conclusions are substantiated by the fact that the trace is symmetrical about B and C, so that the two components must "peak together" at these points.

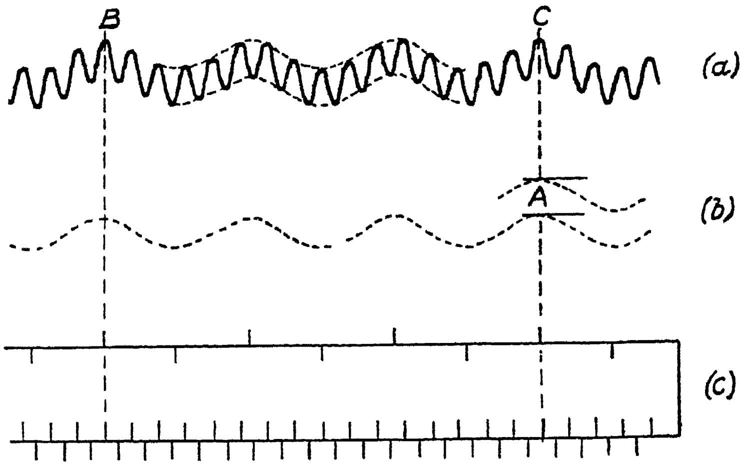


FIG. 4.—Separation of two components of high frequency-ratio. (*b*) shows the low-frequency component of the trace at (*a*), and the double-amplitude of the high-frequency component (A). The short vertical lines in (*c*) represent, in magnitude and position, the peaks of both components. See also Table V, page 115.

The two components can conveniently be represented as shown at (*c*) in Fig. 4. The short vertical lines indicate the crests and troughs, and the length of each line represents the amplitude of the component. Dependent upon the nature of the problem it may or may not be desired to draw the components separately; when it is so desired, the method illustrated offers the advantages of simplicity and the saving of labour. The upper part of (*c*) affords precisely the same information as does the trace at (*b*).

The determination of the cycle of waveforms of this type is easily accomplished by means of the characteristic patterns referred to on page 85. In this way also the existence of a non-harmonic

component may be detected; this component can arise from "hum," induced by the alternating-current mains used in some electrical recording instruments, or from some other source. If the fundamental frequency of the variation recorded (i.e. excluding the hum) is a simple sub-multiple of the mains frequency, the hum will be a harmonic component; but if this condition is not realised, there may be no cycle apparent in the record. Such an absence of cyclic regularity does not invalidate the envelope method of analysis, however; in the case of the type of waveform considered here, the envelopes determine the low-frequency component with equal facility whether the high-frequency component is harmonic or not. On the other hand, slight irregular fluctuations in the variation being recorded may in some cases mask the repetition of the characteristic patterns, so that too great a length of record is taken to

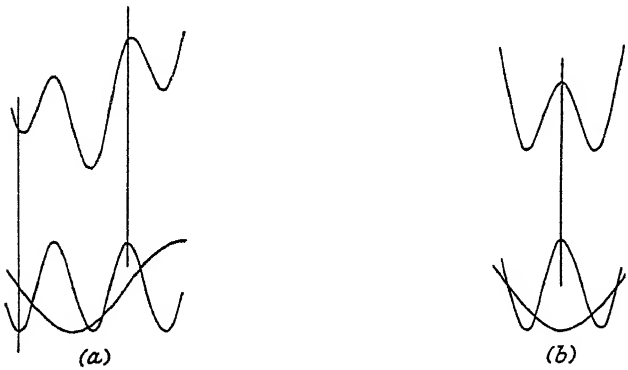


FIG. 5.—(a) Displacement of peaks of ripple when they do not coincide with peaks of low-frequency component; (b) no displacement in the case of coincidence of peaks.

be the cycle. If, for example, two cycles of the wave are taken in error to be the cycle, and the components are really the 2nd and 9th harmonics, they will appear to be the 4th and 18th harmonics of a fundamental frequency which is only half as great as it should be, so that the frequency determinations will not be affected; and the misconception can be detected by the fact that the harmonic numbers have a common factor (2).

The combination of two components of high frequency-ratio is the simplest type of complex wave, so far as analysis is concerned. The process of analysis requires no further explanation, and the reader is advised to practise on the examples illustrated in Fig. 6. It is required to find the harmonic numbers of the two components, to measure their amplitudes and to determine the precise location of one crest or trough of each component (from which the phase

could be calculated with reference to any datum). In the solution of these problems, and similarly in the analysis of any waveform of this type, very effective use can be made of the property of symmetry, illustrated in Fig. 5*b*. By "spotting" configurations of this type the location of peaks is facilitated. Where there appears to be a slight irregularity in the trace, the construction of

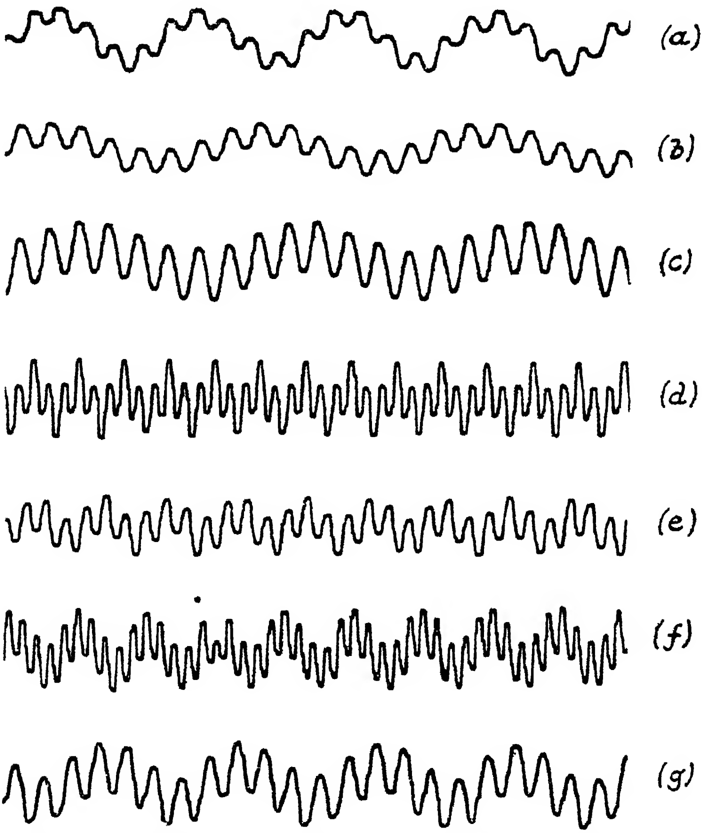


FIG. 6.—Examples of two-component waveforms, for practice.

the envelopes is rendered less ambiguous by the fact that they must be sinusoidal and parallel.

When the reader has analysed several of the examples in Fig. 6, as he is advised to do, the accuracy of the analysis can be checked in the best possible manner—namely, by reconstruction of the traces from a synthesis of the components. The reconstructed trace, if drawn to the same scale as the original, should of course be similar to it. In this manner, and only in this manner, will

confidence be gained; for this reason the "answers" are not supplied. In this connection it should be observed that the traces illustrated have been selected from actual recorded waveforms, and each contains (for all practical purposes) only two components. Traces reconstructed from syntheses of the components may differ in minor respects from the originals, but if the result is different to a major degree the analysis should be repeated to discover the error.

Use of the "Tables for Synthesis" (Table II, Appendix V, page 258) will facilitate these reconstructions.

4. Two-component waves—beating.

The next type of two-component wave to be considered is the beating waveform, resulting from the combination of two components whose frequency difference is small compared with the frequency of either. The synthesis of this type of waveform was described in Chapter I, pages 20, 28, and the following properties should be recalled:

(i) The envelopes are approximately sinusoidal and similar but anti-phased, i.e. crests in one coincide with troughs in the other, and *vice versa.*

(ii) The width of the envelope strip, i.e. the vertical separation of the two envelopes, varies cyclically, the frequency of the variation being the difference between the frequencies of the components.

(iii) The maximum width of the envelope strip equals the sum of the double amplitudes of the components, and the minimum width of the envelope strip equals the difference between these double amplitudes.

(iv) Except in certain critical cases (see p. 23) the frequency of the wave which is visible in the trace is the frequency of the major component (i.e. the component of greater amplitude).

(v) The separation of successive peaks (crests or troughs) at the bulge and at the waist of the beat determines whether the minor component is of higher or lower frequency than the major. If the peak separation at the bulge is greater than at the waist, the frequency of the minor component is less than that of the major; and if the peak separation at the bulge is less than that at the waist, the minor component is of higher frequency than the major.

(vi) The components are temporarily in-phase at the bulge, and temporarily anti-phased at the waist.

These characteristics are exemplified by the waves illustrated in Figs. 10-17, Chapter I.

Fig. 7 shows a beating waveform analysed into its constituent waves. The waveform itself is illustrated at (*a*) and the components at (*b*), where use has been made of the method of depicting waves described on page 89. By considering a greater length of the record (part only of which is shown in the diagram) it has been determined that AA′ is a cycle. Since the trace appears to be symmetrical about the waist B, the components are both cosine waves referred to A or B as datum. The trace is also symmetrical about A, but this symmetry is not so easy to observe in the neighbourhood of a bulge as it is near a waist. Troughs in the two components coincide at A, and a crest of the minor coincides with a trough of the major at B. There are 12 crests (and 12 troughs) in the cycle,. so that the major component is the 12th harmonic.

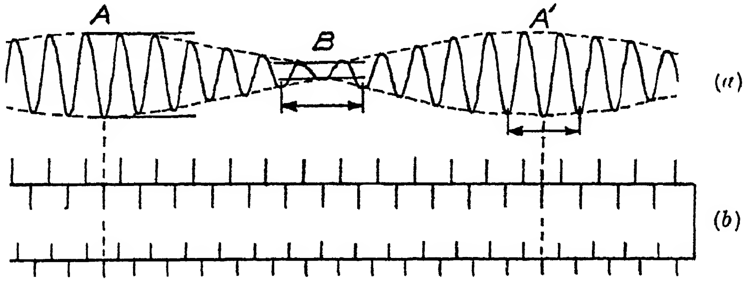


Fig. 7.—Analysis of simple beats.

The peak separation at the waist B is greater than that at the bulge A or A′, as shown by the arrows in the diagram, and according to (v), above, this fact indicates that the minor component is of higher frequency than the major. Finally, there is one beat in the cycle, so that the difference between the harmonic numbers of the components (referred to the cycle) is unity ; hence the minor component is the 13th harmonic.

By means of the envelopes, drawn in broken lines, the amplitudes of the components can be found. In the original trace the maximum strip-width (i.e. the distance between the parallel lines marked at A) was 11·4 mm., and the minimum strip-width (at B) was 2·2 mm. These values are respectively the sum and the difference of the two double amplitudes, which are therefore 6·8 and 4·6 mm. The components are therefore :

12th harmonic : 6·8 mm. (d.a.) * or $\pm$ 3·4 mm.
13th harmonic : 4·6 mm. (d.a.) or $\pm$ 2·3 mm.

* See note on p. 78.

Since it is known that troughs of the two waves coincide at A, the separate components can now be drawn as in the diagram.

When the amplitudes of the two components are equal or nearly equal, so that the minimum strip-width is approximately zero there is occasionally some doubt as to the location of peaks at the waist of the beat. In Fig. 8 the troughs at the waist D may be considered to exist as indicated by the vertical lines below the trace. BC is a cycle of the wave, and assuming the troughs to be as indicated there are 12 cycles of the "major" component. The peak separation at the waist being greater than that at the bulge, the "minor" component is the 13th harmonic; actually, since the minimum strip-width is zero the amplitudes of the two components are equal and the terms "major" and "minor" do not apply. In the cycle AB the location of the troughs at the waist E appears to be as indicated by the vertical lines below that portion of the wave; assuming this location to be correct there are 13 cycles of

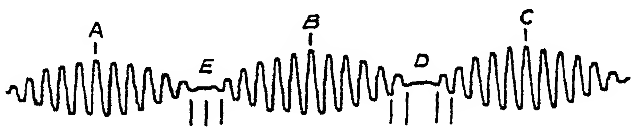


FIG. 8.—Analysis of beats in critical case. It is immaterial whether the location of troughs at the waist is taken to be as marked at D or as marked at E.

the "major" component in the cycle of the beat. Since, however the peak separation at the waist is now less than at the bulge, the "minor" component is the 12th harmonic. Thus it is immaterial which assumption is made; the result in either case is that there are present in the waveform two components, the 12th and the 13th harmonics, of approximately equal amplitude.

The three waveforms in Fig. 9 all have the same fundamental wavelength; the extent of the cycle is shown by the short vertical lines marked above the traces. In all three waves there is one beat per cycle. At (*a*) the major component is the 8th harmonic since there are 8 crests per cycle, and the minor component is therefore either the 7th or the 9th harmonic. By measuring the peak separations at the bulge A and the waist B it is found that the minor component is of lower frequency than the major, so that it is the 7th harmonic. It is sometimes convenient to measure peak separations over a wider range, as at C and D.

The reader may verify for himself that the components of the trace in Fig. 9*b* are the 7th (major) and the 8th (minor) harmonics and that in Fig 9*c* the 7th harmonic is again the major component

he minor component being in this case the 6th harmonic. The)eak separation property is well illustrated by these two waves, which have the same major component.

The trace at (*b*) illustrates also the type of irregular fluctuation eferred to on page 85; the waist C is not quite the same as the waist B. In the complete record, of which the diagram shows only a part, there were twenty beats and the number of crests per beat was constant throughout (seven). Furthermore, the trace is symmetrical about the centre of each waist, so that the characteristic pattern is maintained. Nevertheless, the width of the envelope strip at C is less than at B, and this inequality indicates an irregular fluctuation from cycle to cycle.

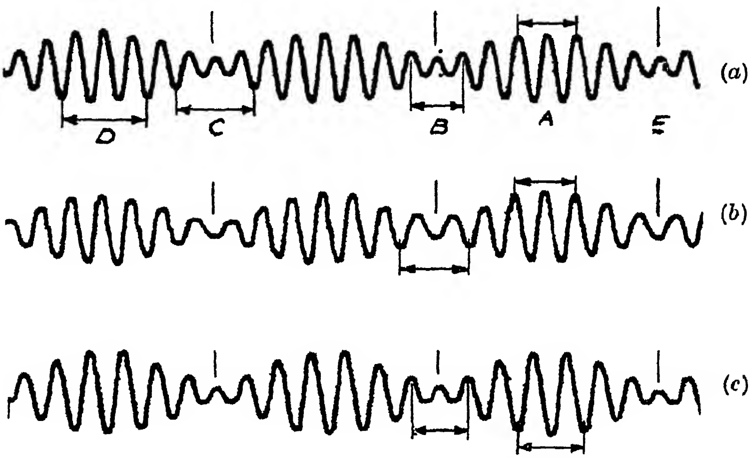


FIG. 9.—Comparison of beats with same beat frequency and a common component.

The trace in Fig. 1, page 86, gives an example of the non-oincidence of beat cycles and fundamental cycles. As already xplained, the cycle of the wave is AA′, which comprises two beats. here are 15 crests in this cycle, so that the components are the h (major) and the 13th or 17th (minor). harmonics. Considera- of peak separation shows the minor component to be the 13th rmonic.

The wave in Fig. 10 has three beats per cycle, the cycle extending from A to A′. It will be observed that the characteristic pattern at A is not repeated at B. Since there are 16 crests in the cycle, and the peak separation at the bulge is greater than that at the waist, the components are the 16th (major) and the 13th (minor) armonics.

Where there is any doubt concerning the repetition or non-repetition of a characteristic pattern at points B between successive definite repetitions A, A′, the following test will be found useful. Compare the number of crests occurring in the portion AA′ with the number of beats in the same interval; if the two numbers possess a common factor other than unity, then that factor is the number of cycles of the complete wave in the interval AA′. Thus in Fig. 10 there are 16 crests and 3 beats in the length AA′; and it is known that AA′ is either one cycle or an integral number of cycles, since the pattern AA′ is repeated regularly throughout the trace (part only of which is shown in the diagram). Since 16 and 3 have no common factor other than unity, the portion AA′ does in fact represent one cycle. On the other hand, the portion CE in Fig. 9 contains two beats and an even number of crests in all three waves, so that this portion comprises two cycles.

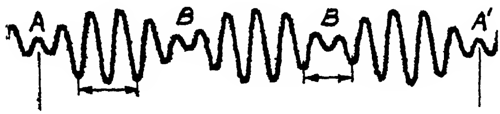


FIG. 10.—Beating waveform with three beats per cycle.

The amplitudes of the two components can in all cases be determined from the maximum and minimum widths of the envelope strip, as explained above in connection with Fig. 7.

5. Phase-determination in beats.

The phases of the components can easily be found from the fact that the two constituent waves are temporarily in-phase at the bulge and anti-phased at the waist; this fact, together with the property of symmetry or skew-symmetry which may be displayed by the trace, completely determines the phases. Thus in Fig. 10, a crest of the 16th harmonic occurs at A, and a trough of the 13th harmonic occurs at the same place; the trace is symmetrical about A, and this fact supports the conclusion. On the other hand, the trace in Fig. 1 is skew-symmetrical about A, so that the two components have zero values there; but the trace is symmetrical about the trough midway between A and B, and both components therefore have troughs there.

In practice it is found that nearly all simple beating waveforms exhibit either symmetry or skew-symmetry, or a condition approaching very nearly to one of these. This fact is a result of the relation between the frequencies of the components, the frequency difference being small compared with either component frequency; at some part of the cycle the two waves are bound to be (either exactly

nearly) in-phase or anti-phased as sine or cosine waves (see Chapter I, p. 11).

It is interesting to compare traces derived from the same components with different phase-relationships. Figs. 11, 12 and 13 show some typical pairs of traces recorded simultaneously. In Fig. 11 the major components of the two traces are in-phase, as shown by the vertical line; but a bulge in one trace coincides with a crest in the other, and *vice versa*. This relation between the beat envelopes is conveniently summarised by the statement that the beats are anti-phased. Since the two components of each trace are temporarily in-phase at a bulge and temporarily anti-phased at a waist, it follows that the minor components of the two traces are anti-phased, so that Fig. 11 illustrates the case:

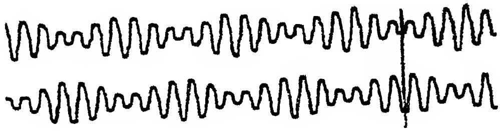


FIG. 11.

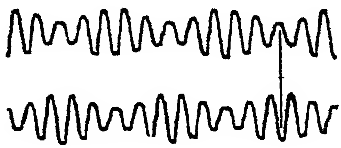


FIG. 12.

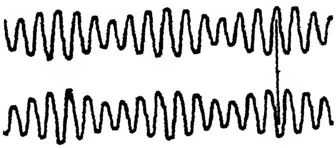


FIG 13.

FIGS. 11-13.—Two beating waveforms recorded simultaneously, showing effects of different phase-relations. The major components are in-phase in Fig. 11 and anti-phased in Figs. 12 and 13; the minor components are in-phase in Fig. 12 and anti-phased in Figs. 11 and 13.

major component : in-phase,
minor component : anti-phased.

In Fig. 12 the major components are anti-phased, as shown by the vertical line, and the beats are anti-phased. Hence the minor components of the two traces are in-phase, and the diagram illustrates the case:

major component : anti-phased,
minor component : in-phase.

Finally, the two traces in Fig. 13 are completely anti-phased, one being the mirror-image (in a horizontal line) of the other; both

components are therefore anti-phased. The same conclusion is reached by the previous method of argument : bulges coincide in the two traces, so that the beats are in-phase, and the major components are anti-phased. Hence Fig. 13 illustrates the case :

major component : anti-phased,
minor component : anti-phased.

These properties are summarised in Table I, where the abbreviations "in" and "anti" denote "in-phase" and "anti-phased," respectively. The major component is the one which contributes the crests and troughs apparent in the traces ; if the phase-relationships of the beats and the major components are known, the phase-relationship of the minor components is found from the third column of the table, and if the phases of the two components are known, the phase-relationship of the beats in the resultant waveforms can be found from the first column of the table.

TABLE I

Beats	Major component	Minor component
In	In	In
In	Anti	Anti
Anti	In	Anti
Anti	Anti	In

6. Other two-component waveforms.

Frequency ratio 2 : 1. Except in special cases, the envelope method of analysis does not yield very accurate results when applied to waveforms whose two components have the frequency ratio 2 : 1. Such waves are best analysed by the method of superposition described in Chapter V. When the 2nd harmonic is reasonably large, a fairly accurate assessment of the harmonic contents can, however, be made by means of a modified form of envelope analysis, which will now be described in connection with a particular example. The wave in Fig. 14 has two components with this frequency ratio. The larger envelope—in this case the top one—is drawn so as to touch the trace at or near the appropriate peaks (in this case the crests). The other envelope is now constructed so as to be as nearly as possible in-phase with the first envelope, and the envelope mean is drawn-in. This line is spaced midway between the two envelopes in the vertical direction (see diagram), and represents part of the fundamental component ; the original

wave is regarded as composed of this wave, together with the components indicated by the beating form of the envelope strip. In the original record from which the diagram is reproduced, the amplitude of the envelope mean was 0·7 in. (d.a.) or $\pm$ 0·35 in.; the maximum width of the envelope strip was 0·9 in. and the minimum width 0·3 in., as shown in the diagram. These measurements are made at the bulge and waist of the beat, which are, of course, separated by a distance equal to half the wavelength. The components of the beat are thus found to be: 2nd harmonic, 0·6 in. (d.a.) or $\pm$ 0·3 in.; fundamental, 0·3 in. (d.a.) or $\pm$ 0·15 in. Since the bulge of the beat, where crests of the two beating components coincide, occurs almost at the same place as the crest of the envelope mean, the total amplitude of the fundamental component is approximately the sum of the amplitudes of the constituent parts, the one represented by the envelope mean and the other by the minor component of the beat. Hence the fundamental is approximately $0{\cdot}7 + 0{\cdot}3 = 1{\cdot}0$ in. (d.a.) or 0·5 in., and the 2nd harmonic has already been completely determined as the major component of the beat. The complete analysis (except for phase-angles) is therefore:

fundamental: 1·0 in. (d.a.) or $\pm$ 0·5 in.
2nd harmonic: 0·6 in. (d.a.) or $\pm$ 0·3 in.

It is a convenient property of this method of analysis that fairly accurate results are obtained even from a very roughly-constructed envelope. In Fig. 15 the lower envelope has been made definitely smaller than in Fig. 14, and the amplitude of the envelope mean was 0·6 in. in the original. The maximum and minimum widths of the envelope strip were 1·02 and 0·23 in., respectively, as indicated in the diagram; the amplitude of the 2nd harmonic is thus found to be $\frac{1}{2}(1{\cdot}02 + 0{\cdot}23) = 0{\cdot}63$ in. (d.a.), and the total amplitude of the fundamental is found to be approximately

$$0{\cdot}6 + \tfrac{1}{2}(1{\cdot}02 - 0{\cdot}23) = 1{\cdot}0 \text{ in.}$$

Comparison of the results obtained from Figs. 14 and 15 indicates the degree of accuracy possible with the method. In cases where the two parts of the fundamental component are not nearly or exactly in-phase, use may be made of the formula (5.2) of Chapter I, page 17. Replacing the phase-difference in this formula by ϕ, the resultant amplitude r is given by

$$r = \sqrt{a^2 + b^2 + 2ab\cos\phi}$$

where a and b are the amplitudes of the two parts of the fundamental component.

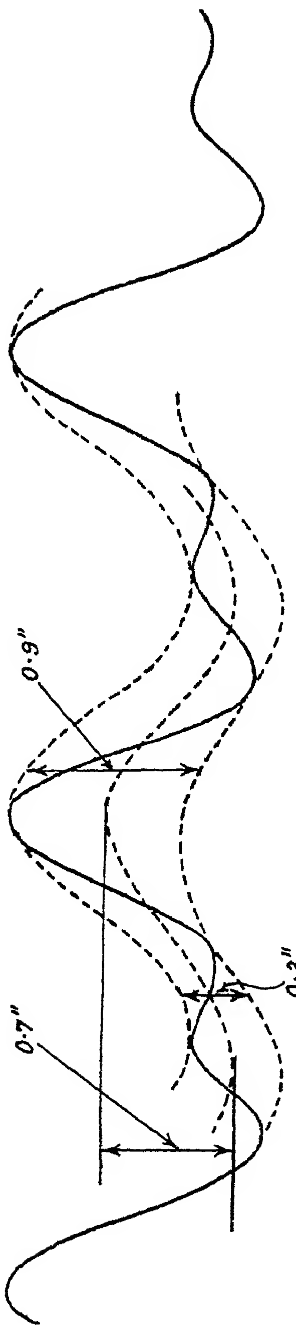


Fig. 14.—Frequency-ratio 2 : 1. Compare Fig. 15.

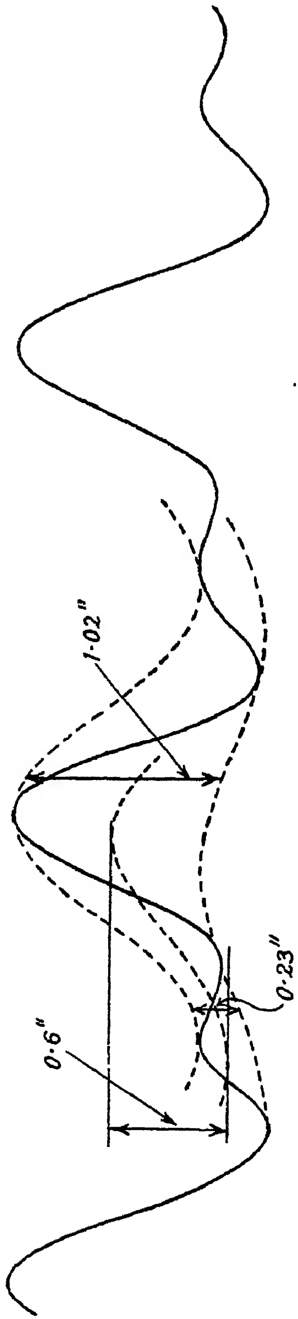


Fig. 15.—Repeat analysis of waveform in Fig. 14, using different envelopes : the final results are the same.

Frequency ratio nearly integral. The trace in Fig. 16 illustrates another type of waveform with two components: the case of a frequency ratio nearly equal to an integer greater than unity. A cycle of the wave extends from A to A′, and by counting it is found that the two components are the 13th and 25th harmonics, so that the frequency ratio is 25 : 13 or nearly 2 : 1. The trace being symmetrical about A, crests of the two components coincide at that place. The small peaks due to the high-frequency component "creep" progressively in relation to the peaks of the low-frequency component, until at B the two components are again temporarily in-phase, with troughs coinciding. It is therefore evident that the distance between the two horizontal lines marked on the diagram (i.e. the overall double amplitude of the trace) is the sum of the double amplitudes of the components. In order to estimate the separate amplitudes, use is made of a special envelope: this is the envelope touching the trace at the *small* troughs, and is shown in

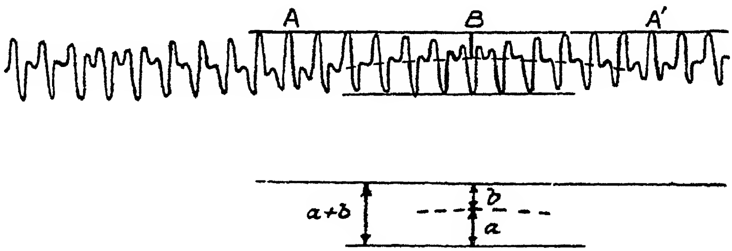


FIG. 16.—Analysis of waveform whose two components have a nearly integral frequency-ratio.

the diagram as a broken line. The distance between this line and the line AA′ at B is the double amplitude of the high-frequency component. The truth of this statement will be evident from a study of the lower part of the diagram, where a, b denote the double amplitudes of the low- and high-frequency components respectively. The broken line at B indicates the position of the small trough that would be there if the low-frequency component had a crest at that point instead of a trough. Other examples of this type can be analysed by the same method.

7. Three-component waveforms.

Waveforms containing three main components can be divided into several general types, according to the relation between the frequencies of the components. Thus, one component may be of very much higher frequency than the other two, appearing as a ripple superimposed on the more gentle variation due to the

lower-frequency components; or one component may be of very much lower frequency than the others, appearing as a low-frequency surge imparted to the trace; finally, the three frequencies may be separated by small amounts, giving a very complex appearance to the wave. The general characteristics of some of these types are first described below, with an outline of the method of analysis, and several examples are then analysed in detail.

Of the various types of waveform with three main components, the easiest to analyse is that wherein a high-frequency "ripple" is superimposed on the sum of the two lower frequency components. The general properties of this type of wave are the same as those discussed in Section 3, except that the low-frequency component now becomes the sum of two low-frequency components. Fig. 17*a*

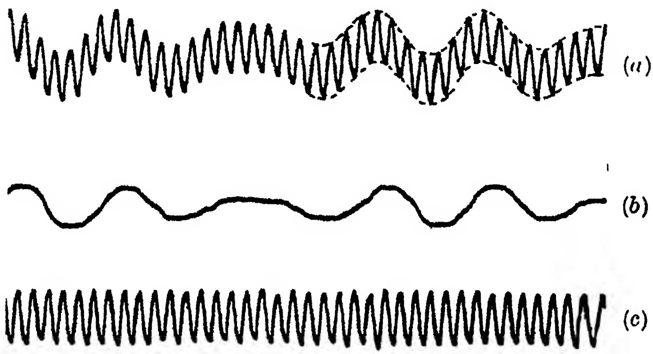


FIG. 17.—Three-component waveform (*a*) with a high-frequency ripple (*c*). The low-frequency components are shown at (*b*).

illustrates an example of this type. The low- and high-frequency components are shown separately at (*b*) and (*c*). [The irregularity observable at (*c*) very frequently appears in actual recorded waveforms.] Since the two envelopes are identical, the envelope mean (or line drawn midway between the envelopes) also represents the low-frequency components but it is not necessary to draw this line. The analysis proceeds in a manner similar to that employed with two-component waves: the vertical separation of the two envelopes gives the double amplitude of the highest frequency component, and the envelopes themselves represent the sum of the two low-frequency components. The envelope can be analysed by the methods described above, and the complete analysis is obtained.

Fig. 18 shows an example of the second type referred to above: one component is of much lower frequency than the other two. AA′ is one cycle of the complete variation (the repetition of the

characteristic pattern should be observed). The high-frequency components are shown at (*b*), where it can be seen that their combination produces one beat per cycle, and the low-frequency component is shown at (*c*). These traces are all actual recorded waveforms, which accounts for slight irregularities. At (*d*) the envelopes are drawn, and also the envelope mean. This line, spaced midway between the envelopes in a vertical direction, represents in all respects the low-frequency component. Thus, for example, the arrows at (*d*) show the double amplitude of the low-frequency surge.

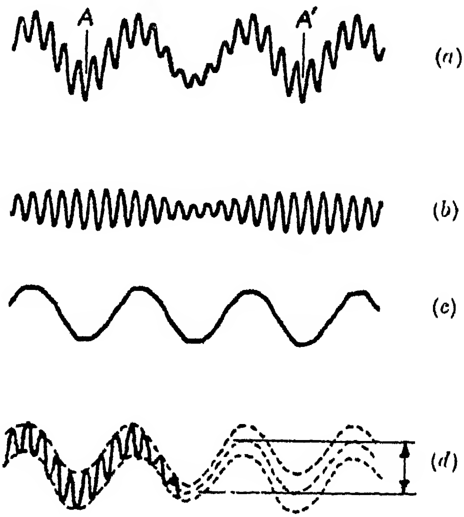


FIG. 18. Three-component waveform (*a*) with a low-frequency surge (*c*). The high-frequency components are shown at (*b*), and the envelopes and envelope mean at (*d*).

A special form of this type of waveform is illustrated in Fig. 19. At (*a*) the trace contains only two components of nearly equal frequencies, which combine to give a beat. The other traces in the

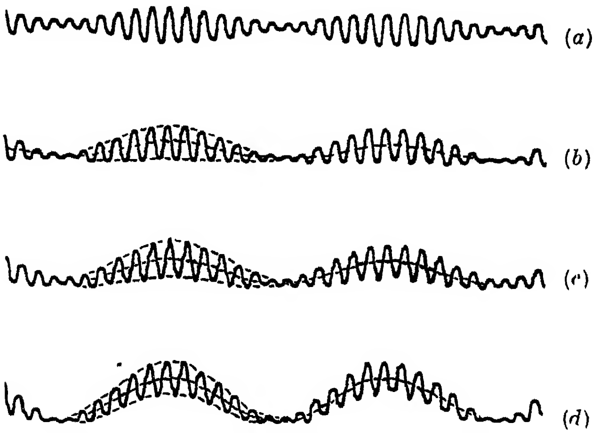


FIG. 19.—The waveforms at (*b*), (*c*) and (*d*) show the result of adding a component at the beat frequency, and in-phase with the top envelope of the simple beat at (*a*).

diagram show the effects of superimposing a wave at the beat frequency, which is in-phase with the top envelope of the beat. At (*b*) the amplitude of this surge is the same as that of the beat envelope, with the result that the bottom envelope at (*b*) is nearly a straight line and the amplitude of the top envelope is double that of the envelope in (*a*). The result is an unsymmetrical beat. At (*c*) and (*d*) low-frequency surges of still greater amplitude are

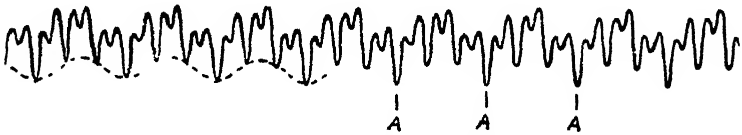


FIG. 20.—Three-component waveform with frequencies in the ratio 6 : 3 : 1.

present. In all cases the envelope mean represents the added low-frequency component. The reader should observe that the envelope mean in (*a*) is a straight line, showing that there is no third component in that trace.

An example of another type of waveform is shown in Fig. 20, where the frequencies are not so widely separated as in Figs. 18 and 19. The higher frequency components have the frequency ratio 2 : 1, and the lines marked A indicate corresponding points in

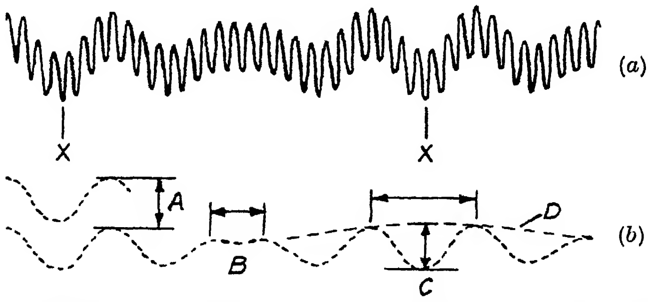


FIG. 21.—Analysis of waveform in Fig. 17*a*, page 102. See also Table VI, page 115.

successive cycles. The low-frequency component is the fundamental, as shown by a modified form of envelope drawn to touch the larger troughs. The components are therefore the 1st, 3rd and 6th harmonics.

Turning now to the detailed analysis of these and other examples, consider first the trace in Fig. 17*a*, reprinted in Fig. 21*a*. XX is a cycle, and there are 24 cycles of the ripple in this interval ; the ripple is therefore the 24th harmonic. The two envelopes being

identical their constant vertical separation (indicated by the arrows at A) is the double amplitude of the ripple. In the original record from which the diagram is reproduced this distance was 6·8 mm. The envelope, which is redrawn at (*b*), represents the low-frequency components. It has a beating form, with a waist at B and a bulge at C, one beat occurring in the cycle. The major component of the beat is the 4th harmonic, and since the peak separation at the bulge C is greater than that at the waist B the minor component is the 3rd harmonic. The three components of the trace at (*a*) are therefore the 3rd, 4th and 24th harmonics. As a check on this result, it may be observed that the harmonic numbers 3, 4 and 24 possess no common factor other than unity.

The line D is the top envelope of the beat, and the vertical arrows at C indicate the sum of the double amplitudes of the 3rd and 4th harmonics. This distance was 6·2 mm. in the original record. The corresponding measurement at the waist B is too small to be shown in the diagram, but was 0·4 mm. in the original record; this is the difference between the double amplitudes of the 3rd and 4th harmonics. These double amplitudes are therefore 3·3 mm. (4th) and 2·9 mm. (3rd). The frequency and amplitude values for all the harmonics have now been determined (the frequencies being given by the harmonic numbers):

3rd harmonic :	2·9 mm. (d.a.) or	1·45 mm.
4th harmonic :	3·3 mm. (d.a.) or	1·65 mm.
24th harmonic :	6·8 mm. (d.a.) or	3·40 mm.

If a timing trace enabled the fundamental period to be calculated, the actual frequencies could be found; thus if the cycle XX occupies a time-interval of 1/10 second, the frequencies are 30, 40 and 240 C.P.S. Alternatively, if the trace records vibrations in an engine running at 1200 R.P.M., the engine orders are $1\frac{1}{2}$E, 2E and 12E, since the frequencies are 1800 ($= 1\frac{1}{2} \times 1200$), 2400 ($= 2 \times 1200$) and 14,400 ($= 12 \times 1200$) C.P.M.

Finally, the phase of each component is given by the fact that all harmonics have a trough at C.

The waveform in Fig. 18*a* is reprinted in Fig. 22. The envelopes and envelope mean are drawn beneath the trace. A and A′ are corresponding points in successive cycles, and it is at once evident that the components are the 2nd, 15th and 14th or 16th harmonics. The peak separation at the bulge B of the beat being slightly greater than that at the waist W, the minor component must be of lower frequency than the major, so that it is the 14th harmonic. The arrows at C and D indicate, respectively, the sum and the difference

of the double amplitudes of the 14th and 15th harmonics, and as these distances in the original record were 5·3 and 1·7 mm., the double amplitudes are found to be 3·5 mm. (15th) and 1·8 mm. (14th). The arrows at E indicate the double amplitude of the 2nd harmonic, which was 7·1 mm. in the original record. The components are therefore :

2nd harmonic :	7·1 mm. (d.a.) or $\pm$ 3·55 mm.
14th harmonic :	1·8 mm. (d.a.) or $\pm$ 0·9 mm.
15th harmonic :	3·5 mm. (d.a.) or $\pm$ 1·75 mm.

Again, all three components have troughs at A, so that the phase-angles can be found with reference to any convenient datum, and the analysis is complete.

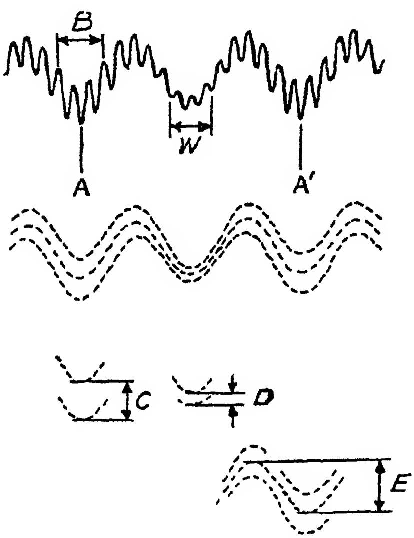


FIG. 22.—Analysis of waveform, in Fig. 18a, page 103. See also Table VII, page 116.

In Fig. 23 the trace shows a waveform representing vibrations in an engine, the recording equipment being electrical in nature. The engine was running at 380 R.P.M., and it is required to measure the amplitudes and frequencies of each component and to determine the source of each. The timing marks in the upper trace indicate intervals of 1/2 second.

PQ is a cycle ; it will be observed that the characteristic pattern at P is repeated at Q and not elsewhere within the range PQ. The ripple is the 79th harmonic of this cycle, and the envelope contains the 5th and 20th harmonics. By means of the timing marks it is found that the frequency of the ripple is 50 C.P.S. or 3000 C.P.M. ; the frequency of the 5th harmonic is therefore $5 \times 3000/79 = 190$ C.P.M., and the frequency of the 20th harmonic is $4 \times 190 = 760$ C.P.M. These two frequencies, being respectively half and twice the engine rotational speed, are identified as the $\frac{1}{2}$E and 2E engine orders, and the ripple is " hum," due to induction from the alternating-current electric mains used to operate the amplifiers in the recording system.

The amplitudes of the various components are easily found in

the usual manner. Thus, in the diagram the arrows at A indicate the double amplitude of the $\frac{1}{2}$E order, those at C or D the double amplitude of the 2E order, and those at B the double amplitude of the hum ripple.

Two points of interest concerning this record should be noted. First, the analysis would not be invalidated if PQ were not a cycle, for the ripple would still be identified as mains hum and the frequencies of the other harmonics must be exact multiples of half the engine speed. Secondly, there is a slight irregularity apparent at X, as shown by the envelope there. This irregularity is due to fluctuations in the recorded vibrations and does not materially affect the analysis.

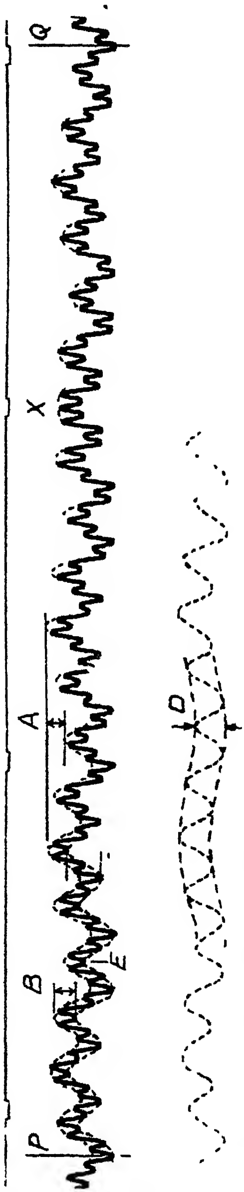


FIG. 23. Mains hum as a high-frequency harmonic. In this wave the mains hum is the 79th harmonic. See Table VIII, page 115.

The waveform in Fig. 24 exemplifies a type of trace, the appearance of which can be misleading. The short, vertical lines below the trace indicate corresponding points in successive cycles. At first sight the components appear to be the fundamental, 2nd and 6th harmonics; for the trace has the appearance of a combination of the 2nd and 6th harmonics (frequency ratio 1 : 3) superimposed upon a fundamental surge of small amplitude. This supposed fundamental component is, however, another example of the apparent low-frequency surge mentioned on page 38. It should be noted that the crests of the ripple at X are much larger than those at Y, which would not be the case if the analysis were as supposed. Construction of the envelopes yields the true result: the envelope mean consists of the 2nd harmonic alone, as shown in the diagram, and the 6th harmonic is beating with either the 5th or the 7th, B and W in the diagram indicating the bulge and waist of the beat. The peak

separation at the bulge B is greater than that at the waist W, as shown, so that the minor component of the beat is the 5th harmonic. The amplitudes of the components can be found in the usual manner from measurements of the strip-width and the envelope mean.

It will be seen that care has to be exercised in applying the envelope method of analysis; in the case of the waveform in Fig. 24 the presence or absence of a fundamental component might have been of considerable importance. The construction of the envelopes and the envelope mean normally gives, as here, a fairly accurate indication of the lower-frequency components and so proves in this case the absence of the fundamental; the danger lies, however, in over-confidence in the analyst. After the routine analysis of many hundreds of records by the envelope method, a certain facility is acquired which may lead either to increased or

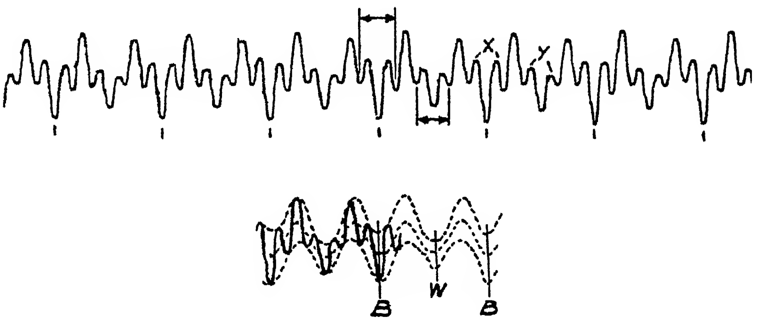


FIG. 24.—Apparent low-frequency surge due to beating of high-frequency components. The true components are the 2nd, 5th and 6th harmonics, and there is no fundamental component.

to decreased accuracy, according to the individual. It is for this reason that the early part of the present volume describes the synthesis of complex waves from their components; a thorough understanding and wide experience of the way in which sine-waves of different frequencies combine is of inestimable value to the would-be analyst, as they will prevent him from "jumping to conclusions" too readily.

In Figs. 25-27 the envelopes are drawn, and some of the additional constructions or measurements shown, and the complete analysis of these traces, and of that in Fig. 2, p. 87, is left to the reader as an exercise. Brief notes, including the correct harmonic numbers of the components, are given below, and the reader is urged to perform the analysis himself, measuring the amplitudes of the various components and determining as accurately as possible

the phase angles with reference to some convenient datum, finally checking the results by a synthesis of the components.

Fig. 25. AB represents two cycles, despite the slight irregularity discernible. It is immaterial whether the peak separations in the beat are taken to be as indicated at B, C or as at A, D. The components are the 4th, 5th and 15th harmonics.

Fig. 26. The cycle is clearly defined in the trace. The components are the 3rd, 12th and 13th harmonics.

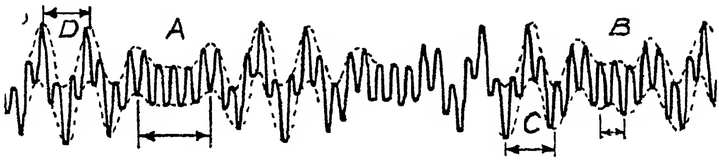


FIG. 25.

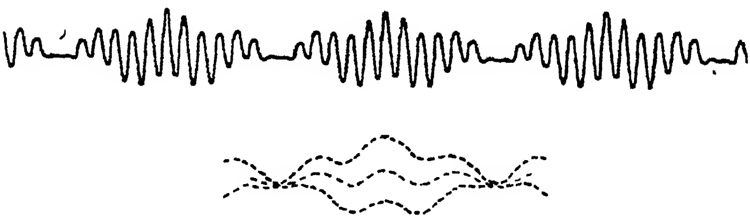


FIG. 26.

FIG. 27.

FIGS. 25-27.- Examples for practice; the harmonic numbers of the components are given in the Text.

Fig. 27. The lines marked A indicate corresponding points in successive cycles. The components are the 2nd, 5th and 6th harmonics.

Fig. 2 (p. 87). AB is a cycle. The components are the 2nd, 12th and 13th harmonics.

8. General plan of attack; systematic treatment.

The majority of waveforms obtained in the recording of mechanical vibrations contain at the most three components, and the frequencies of these components are usually so related that

the method of analysis described in the foregoing sections may be used. The general plan of attack may be summarised as follows.

(i) *Determination of cycle.*

Wherever possible the wavelength of the record should be determined. Use may be made, if necessary, of the characteristic patterns.

The absence of a well-defined cycle does not preclude the possibility of using the envelope method.

It is better to take too great a distance for the wavelength than too short a distance, so that in cases where there is doubt the longer distance should be taken.

If the presence of an extraneous component is suspected—mains hum, for example—it may be evident that if this component were absent from the record the cycle would be well defined. In such cases the effects of the extraneous component can be disregarded in the determination of the cycle for the remaining components.

(ii) *Separation of low- and high-frequency components.*

Construction of the envelopes and the envelope mean enable the low- and high-frequency components to be separated.

If the two envelopes are identical there is no necessity to construct the envelope mean, which would have the same form as the envelopes. In this case there is a single high-frequency component.

In all cases the envelope mean represents the low-frequency component or components.

In the construction of the envelopes, care must be taken to include all the peaks of the highest frequency wave observable in the trace, even when these peaks are masked in some parts of the record. Where the spacing of the peaks directly observable in the trace is too irregular to be accounted for by beat properties, peaks must be inserted so as to preserve a more-or-less even distribution.

If the two envelopes are not identical it will usually be found that the width of the envelope strip—i.e. the area between the two envelopes—varies cyclically, indicating a beat in the high-frequency components.

(iii) *Determination of frequencies.*

The fundamental frequency—i.e. the reciprocal of the period of time occupied by one cycle of the complex wave—is found from supplementary timing marks or by reference to special frequency standards incorporated in the record.

As a last resort the frequency can be calculated from the measured film speed, i.e. the speed at which the film or paper bearing the record moves through the recording instrument.

The frequencies of the various components are found by multiplying their harmonic numbers by the fundamental frequency.

Determination of harmonic numbers. The harmonic number of a high-frequency ripple, a low-frequency surge or the major component of a beat may be found by counting the number of cycles thereof that occur in the fundamental cycle.

The number of beats occurring in the fundamental cycle is the difference between the harmonic numbers of the major and minor components.

The minor component of a beat is of higher or lower frequency than the major component according as the peak separation at the bulge of the beat is less or greater than the peak separation at the waist of the beat.

Non-harmonic components. The frequency of a non-harmonic component—i.e. one whose source is distinct from that of the other components, and which is not a harmonic of the main variation—is found by reference to the timing marks or other frequency standard. The fundamental cycle of the main variation may be taken as a frequency standard if convenient.

(iv) *Determination of amplitudes.*

The double amplitude of a single high-frequency ripple is the (constant) width of the envelope strip—i.e. the vertical separation of the two envelopes.

The amplitude of a single low-frequency surge is found from the envelope mean, which represents the surge in all respects.

In a beating waveform, the strip-width at the bulge is the sum of the double amplitudes of the components, and the strip-width at the waist is the difference between these double amplitudes.

In other combinations of two components, the amplitudes can be determined if two points A and B can be found, such that at A crests or troughs of both components coincide, while at B a crest of one coincides with a trough of the other.

Use may also be made of special modified envelopes, as in Figs. 16 and 20.

(v) *Determination of phase-angles.*

With reference to any convenient datum, the phase-angle of a wave can be found if any crest or trough can be located.

Location of peaks. In beating waveforms the components are temporarily in-phase at the bulge and temporarily anti-phased at the waist.

TABLE

Observe distribution of peaks apparent in wave and	
Construct envelopes to touch (i) at or near all the crests,	
If envelopes identical and in-phase, there is a single H.F. component.	
L.F.	H.F.
represented by either envelope. Use as waveform for separate analysis by this table, unless it is a single sine-wave.	"fills-in" the space between the envelopes.
When a single component is obtained, analyse by	

TABLE

	(*a*) Single component.	
	From envelopes.	H.F. ripple.
Harmonic number.	Number of crests per cycle.	
Double amplitude.	Overall amplitude of wave.	Distance between the two envelopes.
Location of peaks.	As in wave.	As at those parts of wave where L.F. component has peaks.
	Use properties of symmetry and skew-	
	If wave is symmetrical about point P, all components have peaks at P.	

II

indicate any that may be masked or otherwise concealed.
(ii) at or near all the troughs, whether apparent or indicated.

If envelopes not identical, or not in-phase, construct envelope mean.	
L.F.	H.F.
represented by envelope mean. Use as waveform for separate analysis by this table, unless it is a single sine-wave.	beating. Consider the variation of distance between the two envelopes, determining bulges and waists, measure peak separations at bulge and waist, and count the number of beats per cycle.
Table III (*a*).	Analyse by Table III (*b*).

III

(*b*) Beating.		
Major component.	Minor component.	
Number of crests per cycle.	If peak separation at bulge is less than at waist, sum of harmonic number of major component and number of beats per cycle.	If peak separation at bulge is greater than at waist, difference between number of major component and number of beats per cycle.
Half the sum of the maximum and minimum distances between the envelopes.	Half the difference between the maximum and minimum distances between the envelopes.	
As indicated at centre of bulges and waists.	In-phase with major component at bulge, and anti-phased with major component at waist.	

symmetry where possible :
If wave is skew-symmetrical about point P, all components have zero values at P.

If the trace is symmetrical about any point, all the components have peaks (crests or troughs) at that point.

If the trace is skew-symmetrical about any point, all the components have zero values at that point.

If the trace is symmetrical or skew-symmetrical about any point in the immediate neighbourhood of that point, the above conclusions regarding peak or zero-point location hold with reference to the higher harmonics.

The location of peaks in a ripple is most accurate where the low-frequency component has a peak, since there is no horizontal displacement of peaks of the ripple at such points.

The procedure of envelope analysis is summarised in Tables II and III, where "L.F." and "H.F." denote "low frequency" and "high frequency" respectively, and "cycle" refers to the cycle of the complete wave. The tables apply to normal types of waveform where no two components have a frequency-ratio 2 : 1 or nearly equal to an integer greater than unity.

Tables IV-IX show the method of applying the general Tables II and III to the particular cases illustrated in Figs. 1, 4, 21, 22, 23, 28. Only the relevant parts of Tables II and III are quoted in each case, dependent upon the nature of the components as indicated by the systematic method of reduction.

It will be noted that the waveform in Fig. 28, analysed in Table IX, contains four main components. The general Tables II and III are capable of handling such waves as easily as they do

TABLE IV

Analysis of waveform in Fig. 1 (p. 86)

<table>
<tr><td colspan="3">Envelopes neither identical nor in-phase; construct envelope mean</td></tr>
<tr><td>Envelope mean is straight line; hence no L.F. component.</td><td colspan="2">Beating. Peak separation at bulge greater than at waist; 2 beats per cycle. Continue as under.</td></tr>
<tr><td></td><td>Major component.</td><td>Minor component.</td></tr>
<tr><td>Harmonic number.</td><td>15</td><td>15 − 2 = 13</td></tr>
<tr><td>Double amplitude.</td><td colspan="2">As in Table III.</td></tr>
<tr><td>Location of peaks.</td><td colspan="2">Symmetry shows that troughs of both components occur midway between waists A and B.</td></tr>
</table>

TABLE V

Analysis of waveform in Fig. 4 (p. 89)

<table>
<tr><td colspan="3">Envelopes identical and in-phase.</td></tr>
<tr><td colspan="2">Envelope represents L.F. component, which is a single sine-wave and is immediately analysable.</td><td>H.F. ripple.</td></tr>
<tr><td>Harmonic number</td><td>(L.F.) 3</td><td>(H.F.) 16</td></tr>
<tr><td>Double amplitude.</td><td colspan="2">As in Table III.</td></tr>
<tr><td>Location of peaks.</td><td colspan="2">Symmetry shows that crests of both components occur at B.</td></tr>
</table>

TABLE VI

Analysis of waveform in Fig. 21 (p. 104)

<table>
<tr><td colspan="4">Envelopes identical and in-phase.</td></tr>
<tr><td colspan="3">Envelope represents L.F. components, which form a complex wave. Analyse this separately as under.</td><td rowspan="3">H.F. ripple. Continue as under.</td></tr>
<tr><td colspan="3">Envelopes neither identical nor in-phase. Construct envelope mean.</td></tr>
<tr><td>Envelope mean is straight line. Hence no component other than those of beat.</td><td colspan="2">Beating. Peak separation at bulge greater than at waist; one beat per cycle.</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">H.F. ripple.</td><td colspan="2">Beats.</td></tr>
<tr><td>Major component.</td><td>Minor component.</td></tr>
<tr><td>Harmonic number.</td><td>24</td><td>4</td><td>$4 - 1 = 3$</td></tr>
<tr><td>Double amplitude</td><td>A in diagram.</td><td colspan="2">As in Table III.</td></tr>
<tr><td>Location of peaks.</td><td colspan="3">Symmetry shows all components to have troughs at X</td></tr>
</table>

TABLE VII

Analysis of waveform in Fig. 22 (p. 106)

<table>
<tr><td colspan="4">Envelopes not identical; construct envelope mean.</td></tr>
<tr><td colspan="2">Envelope mean is a sine-wave representing L.F. component. Continue as under.</td><td colspan="2">H.F. components beating. Peak separation at bulge greater than at waist; one beat per cycle.</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">L.F.</td><td colspan="2">Beats.</td></tr>
<tr><td>Major component.</td><td>Minor component.</td></tr>
<tr><td>Harmonic number</td><td>2</td><td>15</td><td>15 − 1 = 14</td></tr>
<tr><td>Double amplitude</td><td>E in diagram.</td><td>$\frac{1}{2}$(C + D) in diagram.</td><td>$\frac{1}{2}$(C − D) in diagram.</td></tr>
<tr><td>Location of peaks.</td><td colspan="3">Symmetry shows all components to have troughs at A.</td></tr>
</table>

TABLE VIII

Analysis of waveform in Fig. 23 (p. 107)

<table>
<tr><td colspan="4">Envelopes identical and in-phase.</td></tr>
<tr><td colspan="3">Envelopes represent L.F. component, which form a complex wave. Analyse this separately as under.</td><td rowspan="3">H.F. ripple. Continue as under.</td></tr>
<tr><td colspan="3">Envelopes identical and in-phase.</td></tr>
<tr><td colspan="2">Envelope is sine-wave, representing L.F. component.</td><td>Medium-frequency ripple. Continue as under.</td></tr>
<tr><td></td><td>L.F.</td><td>Medium F.</td><td>H.F.</td></tr>
<tr><td>Harmonic number</td><td>5</td><td>20</td><td>79</td></tr>
<tr><td>Double amplitude</td><td>A in diagram.</td><td>C or D in diagram.</td><td>B in diagram.</td></tr>
<tr><td>Location of peaks.</td><td>Trough at D.</td><td>Crest at D.</td><td>Symmetry in neighbourhood of E shows ripple to have trough thereat.</td></tr>
</table>

TABLE IX

Analysis of waveform in Fig. 28 (p. 118)

<table>
<tr><td colspan="5">Envelopes not identical; construct envelope mean.</td></tr>
<tr><td colspan="3">Envelope mean represents L.F. components, which form complex wave. Analyse this separately as under.</td><td colspan="2" rowspan="3">H.F. components beating; peak separation at bulge greater than at waist (not evident in diagram, but clear in the original waveform); one beat per cycle.</td></tr>
<tr><td colspan="3">Envelopes identical and in-phase.</td></tr>
<tr><td colspan="2">Envelopes represent L.F. component which is a single sine-wave.</td><td>Medium-frequency ripple. Continue as under.</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">L.F.</td><td rowspan="2">Medium F.</td><td colspan="2">Beats.</td></tr>
<tr><td>Major cpt.</td><td>Minor cpt.</td></tr>
<tr><td>Harmonic number</td><td>1</td><td>4</td><td>15</td><td>15 − 1 = 14</td></tr>
<tr><td>Double amplitude.</td><td>D in diagram.</td><td>C in diagram.</td><td>½(B + W) in diagram.</td><td>½(B − W) in diagram.</td></tr>
<tr><td>Location of peaks.</td><td>Crest at C.</td><td>Crest at C.</td><td colspan="2">Cannot be found accurately; probably both components have crests at B.</td></tr>
</table>

three-component waves, with the same restrictions on the relations between the frequencies of the components. A certain loss of accuracy is bound to result, however, from the drawing of numerous envelopes for the same trace, and too much reliance should not be placed on the amplitude figures obtained in such cases, unless the waveform is greatly magnified before analysis.

The course followed in the systematic scheme of reduction is indicated in Fig. 29.

The reader is advised to complete tables similar to Tables IV-IX for the other waveforms illustrated in this chapter, and to use a similar scheme of reduction for routine analysis of waveforms that can be handled by the envelope method.

9. Practical notes.

The degree of accuracy obtainable in envelope analysis is enhanced by magnification of the trace. This magnification can

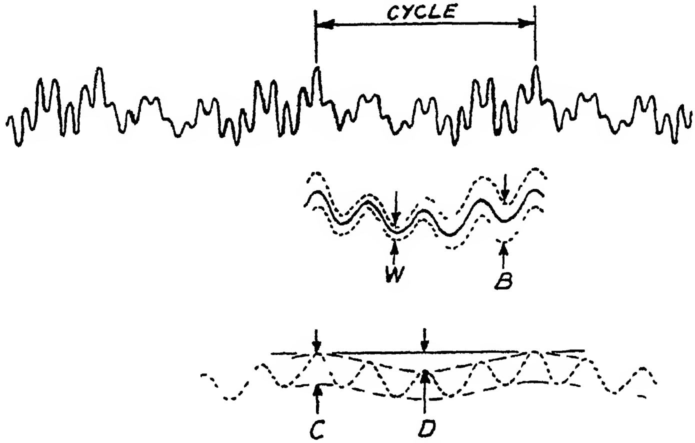


FIG. 28.—Waveform with four components, analysed in Table IX.

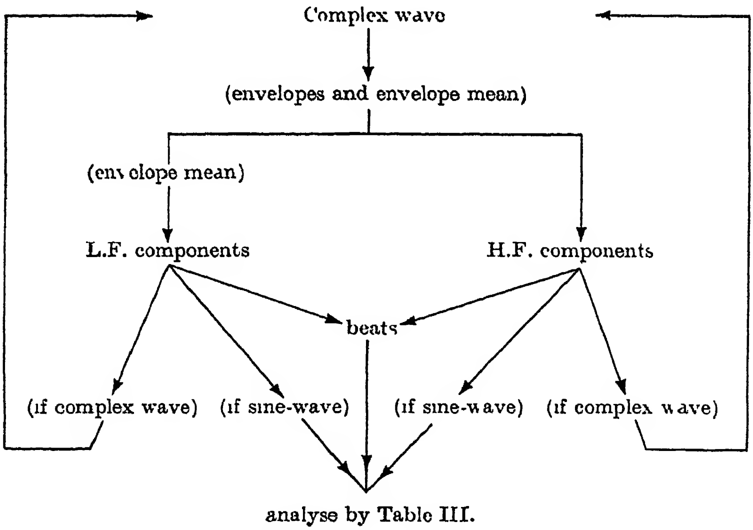


FIG. 29.—Schematic representation of the general procedure in the envelope method of analysis.

be achieved by direct optical projection or by photographic reproduction. Care should, of course, be taken in either case to include all time-reference markings and similar indications in the enlargement.

The use of parallel rules, made in some suitable transparent material, simplifies the determination of overall amplitudes of complex traces, and various other amplitude measurements. It is a fairly simple matter to construct a parallel rule operated by means of a cam, the cam being turned by means of a knob with a dial attached, which is calibrated so as to give a direct indication either of amplitudes or of double amplitudes, according to requirements. Alternatively, the cam may be coupled to a potentiometer which is connected into an electrical resistance bridge circuit; variable resistances in two of the other arms of the bridge are set so as to represent calibration constants, so that the setting of a variable resistance in the fourth arm to give a balanced bridge indicates some desired physical quantity related to the amplitude measured with the rule.

The construction of envelopes is best performed on slips of tracing paper. The process described on page 129 can be applied to the construction of the envelope mean, using the envelopes as data. Tracing paper is also useful for determining the cycle in complicated waveforms; a length of record is traced-off, and the tracing is then applied to other parts of the record, to discover where repetition occurs.

The general notes in Chapter III concerning measurement of amplitudes, frequencies and phase-angles will, of course, be kept in mind during all analysis.

CHAPTER V

METHOD OF SUPERPOSITION

1. Introductory.

The process of analysing a complex waveform into its component sine-waves has been described in the preceding chapters, where the complex waves considered have been of such a type as to enable the analysis to be performed by means of the envelope method. These waves have in general possessed two or three predominant components. There are, however, many complex waves which do not fall into this category and which require analysis. For the complete determination of the components of such waves mathematical or numerical methods must usually be employed, and these methods are described in succeeding chapters; but there is a process whereby a complex wave may be systematically simplified, by separating the odd harmonics from the even harmonics, or the sine components from the cosine components, or the 3rd, 6th, 9th, etc., harmonics from the remainder, and so on.

This process is termed the "method of superposition," and consists in principle of dividing the cycle into a number of equal sub-cycles and determining the variation which is the average of the variations in the different sub-cycles. A brief description of the method, with some examples, is given by Humphrey, reference 1 in the Bibliography at the end of the book.

The present chapter describes the theory of the method, and illustrates it by application to a waveform which cannot be analysed by the inspection method; the application to waveforms with two components of frequency ratio 2 : 1 or 3 : 1, or with three components with simple integral frequency ratios such as 3 : 2 : 1, is then illustrated, and the main part of the chapter concludes with an indication of the method whereby in most cases the harmonics of a waveform containing components up to the 12th harmonic can be separated out with the minimum of computation.

The final section gives the proof of a general mathematical theorem, the result of which is utilised in establishing the theoretical basis of the method.

2. Theory of the method.

A periodic variation y can be expressed in terms of the independent variable x in the form (4.1) of Chapter II, page 49, i.e.

$$y = f(x) = A_0 + \sum_{1} A_k \sin (kx + \phi_k). \quad . \quad . \quad (2.1)$$

This variation has a period 2π in x. If the basic variable x is increased by π, i.e. by a half-period, the modified function y' is given by

$$y' = f(x + \pi) = A_0 + \sum_k A_k \sin [k(x + \pi) + \phi_k]. \qquad (2.2)$$

The effect of the half-period shift is to change the sign of the odd harmonics, as seen in Section 5 (ii), page 52 ; if N takes all integral values, including unity, the equation (2.2) can be written as

$$y' = f(x + \pi) = A_0 + \sum_N A_{2N} \sin [2N(x + \pi) + \phi_{2N}]$$
$$+ \sum_N A_{2N-1} \sin [(2N - 1)(x + \pi) + \phi_{2N-1}],$$

and since $\quad \sin [2N(x + \pi) + \phi_{2N}] = \sin [(2N)x + \phi_{2N}],$

and $\sin [(2N - 1)(x + \pi) + \phi_{2N-1}] = - \sin [(2N - 1)x + \phi_{2N-1}],$

$$y' = f(x + \pi) = A_0 + \sum_N A_{2N} \sin [(2N)x + \phi_{2N}]$$
$$- \sum_N A_{2N-1} \sin [(2N - 1)x + \phi_{2N-1}]. \qquad (2.3)$$

From (2.1) and (2.3), the addition of y and y' yields

$$y + y' = f(x) + f(x + \pi) = 2\{A_0 + \sum_N A_{2N} \sin [(2N)x + \phi_{2N}]\}, \qquad (2.4a)$$

while the subtraction of y' from y yields

$$y - y' = f(x) - f(x + \pi) = 2 \sum_N A_{2N-1} \sin [(2N - 1)x + \phi_{2N-1}]. \qquad (2.4b)$$

The expressions on the right-hand side of the equations (2.4) are, respectively, twice the even harmonics, and twice the odd harmonics, of the original variation (2.1), so that by adding and subtracting two successive half-cycles of the function the even and odd harmonics are separated. Before considering the immediate application of this method of analysis, it will be convenient to determine whether any extension of the method is practicable.

For example, in the same manner as the complete cycle has been divided into two sub-cycles, subsequently added together to yield twice the even harmonics, it is reasonable to suppose that if the cycle is divided into three equal parts which are subsequently added together, the result might yield three times the sum of those harmonics whose reference number is a multiple of three—i.e. the 3rd, 6th, 9th, etc., harmonics. Investigation shows this to be in fact the case : let the harmonic reference number $k = 3N$, $3N - 1$ or $3N - 2$, where N is integral, according as the remainder from the

division of k by 3 is 0, 2 or 1. Then

$$y = f(x) = A_0 + \sum_N A_{3N} \sin [(3N)x + \phi_{3N}]$$
$$+ \sum_N A_{3N-1} \sin [(3N-1)x + \phi_{3N-1}]$$
$$+ \sum_N A_{3N-2} \sin [(3N-2)x + \phi_{3N-2}]. \quad . \quad . \quad (2.5)$$

The function obtained by increasing the basic variable x by one-third of a cycle, i.e. by $2\pi/3$, is

$$f(x + 2\pi/3) = A_0 + \sum_N A_{3N} \sin [(3N)x + \phi_{3N}]$$
$$+ \sum_N A_{3N-1} \sin [(3N-1)x - 2\pi/3 + \phi_{3N-1}]$$
$$+ \sum_N A_{3N-2} \sin [(3N-2)x - 4\pi/3 + \phi_{3N-2}], \quad (2.6)$$

since $\sin (\theta + 2\pi) = \sin \theta$, while the function obtained by increasing x by two-thirds of a cycle, i.e. by $4\pi/3$, is

$$f(x + 4\pi/3) = A_0 + \sum_N A_{3N} \sin [(3N)x + \phi_{3N}]$$
$$+ \sum_N A_{3N-1} \sin [(3N-1)x - 4\pi/3 + \phi_{3N-1}]$$
$$+ \sum_N A_{3N-2} \sin [(3N-2)x - 2\pi/3 + \phi_{3N-2}]. \quad (2.7)$$

Now $$\sin \theta + \sin (\theta - 2\pi/3) + \sin (\theta - 4\pi/3) = 0,$$

(for a proof of this result see Section 6, p. 135), so that when the three functions (2.5, 6, 7) are added together,

$$f(x) + f(x + 2\pi/3) + f(x + 4\pi/3)$$
$$= 3\{A_0 + \sum_N A_{3N} \sin [(3N)x + \phi_{3N}]\} \quad (2.8)$$

The process of adding together the three parts of the complete cycle therefore results in the separation of those harmonics whose reference numbers are multiples of 3.

In a similar manner, if the cycle is divided into n equal parts, where n is an integer, the functions are

$$f(x),\ f(x + 2\pi/n),\ f(x + 4\pi/n),\ \ldots\ f[x + (n-1)(2\pi/n)],$$

and when these are all added together the result is

$$n\{A_0 + \sum_N A_{nN} \sin [(nN)x + \phi_{nN}]\}. \quad . \quad . \quad (2.9)$$

This result depends upon the general formula of which the particular case $n = 3$ has been quoted above, viz.:

$$\sin\theta + \sin(\theta - 2\pi/n) + \sin(\theta - 4\pi/n) \ldots \sin[\theta - (n-1)(2\pi/n)] = 0.$$

A proof of this formula is given on page 135.

It should be noted that the constant term A_0 appears in the general formula (2.9), since zero may be regarded as a multiple of any integer.

The results obtained above are independent of the choice of a datum-line for the measurement of the basic variable x, i.e. it does not matter at what part of the variation the cycle is supposed to commence. The only effect of a change in the datum-line for measurement of x will be to alter all the phase-angles of the harmonics.

If the variation (2.1) is expressed in the alternative form (4.2) of Chapter II, page 49, i.e.

$$y = f(x) = A_0 + \sum_k a_k \cos kx + \sum_k b_k \sin kx, \quad . \quad (2.10)$$

the general result (2.9) can similarly be transcribed in the form

$$n\{A_0 + \sum_N a_{nN} \cos(nN)x + \sum_N b_{nN} \sin(nN)x\}. \quad . \quad (2.11)$$

A further procedure of the same general type consists of reversing the wave along the axis of the independent variable, i.e. changing the sign of x:

$$y' = f(-x) = A_0 + \sum_k a_k \cos kx - \sum_k b_k \sin kx, \quad (2.12)$$

since $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$. Adding y and y',

$$f(x) + f(-x) = 2[A_0 + \sum_k a_k \cos kx], \quad . \quad (2.13a)$$

and subtracting y' from y,

$$f(x) - f(-x) = 2\sum_k b_k \sin kx. \quad . \quad . \quad (2.13b)$$

The formulæ (2.13) therefore enable the sine and cosine components of any periodic wave to be separated.

3. Example of application to analysis.

Suppose that a periodic variation y is given in the form of twenty-four values spaced at equal intervals over the cycle, $O - 2\pi$ in x, the values being:

x°.	y	x°.	y	x°	y	x°	y
0	39·4	90	12·6	180	13·9	270	27·9
15	29·4	105	6·3	195	10·9	285	19·5
30	26·5	120	17·7	210	19·5	300	16·3
45	26·5	135	27·6	225	23·7	315	16·1
60	17·1	150	18·3	240	15·6	330	15·2
75	15·0	165	14·0	255	21·0	345	30·0

(3·1)

These values might alternatively have been obtained from measurements of a recorded graph of the variation (Fig. 1*a*). The graph is not symmetrical, skew-symmetrical or alternant, so that both sine and cosine components may be expected, as also both even and odd harmonics.

In Table I the first twelve values, from $x = 0$ to $x = 165°$, are listed in order under (*a*), and the last twelve values, from $x = 180°$ to $x = 345°$, are listed opposite them in the column under (*b*). The sum of (*a*) and (*b*) is given under (*c*), and this column represents

TABLE I

x°.		(*a*).	(*b*).	(*c*) = (*a*) + (*b*).	$d = \frac{1}{2}(c)$.	(*e*) = (*a*) − (*b*).	$f = \frac{1}{2}(e)$.
0	180	39·4	13·9	53·3	26·7	25·5	12·8
15	195	29·4	10·9	40·3	20·2	18·5	9·3
30	210	26·5	19·5	46·0	23·0	7·0	3·5
45	225	26·5	23·7	50·2	25·1	2·8	1·4
60	240	17·1	15·6	32·7	16·4	1·5	0·8
75	255	15·0	21·0	36·0	18·0	− 6·0	− 3·0
90	270	12·6	27·9	40·5	20·3	−17·3	− 8·7
105	285	6·3	19·5	25·8	12·9	13·2	− 6·6
120	300	17·7	16·3	34·0	17·0	1·4	0·7
135	315	27·6	16·1	43·7	21·9	11·5	5·8
150	330	18·3	15·2	33·5	16·8	3·1	1·6
165	345	14·0	30·0	44·0	22·0	−16·0	− 8·0
					26·7		−12·8
					etc.		etc.

twice the sum of the even harmonics; similarly, the difference (*a*) − (*b*) is given under (*e*), and this column represents twice the sum of the odd harmonics. Columns (*d*) and (*f*) are respectively half of (*c*) and (*e*) (correct to one decimal place), and the corresponding points are plotted in Fig. 1*b*, *c*. The values given in the table refer, of course, only to half the original cycle, from 0 to 165°;

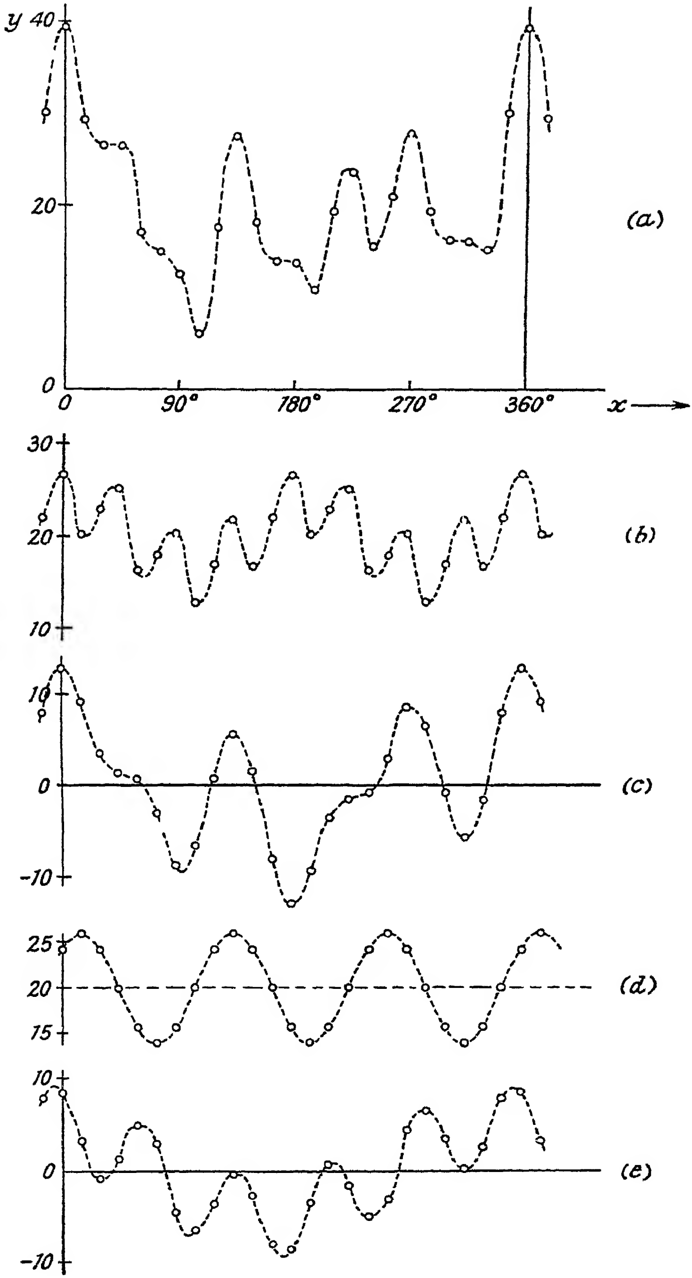


Fig. 1.—Analysis by superposition method. (*a*) Waveform to be analysed, defined by the points marked with circles; (*b*) even harmonics, Table I(*d*); (*c*) odd harmonics, Table I(*f*); (*d*) third harmonic, Table II(*e*); (*e*) first and fifth harmonics, Table III (*c*).

in the case of the even harmonics the second half-cycle is identical with the first, since the period of the even harmonic series is 180° in x; and in the case of the odd harmonics the series is alternant, so that the values for x from 180° to 345° are the same as those for the first half-cycle but with the signs reversed.

The diagram (*b*) is a waveform whose components are clearly the 2nd and 8th harmonics, and it is easily analysed by the envelope method as a wave containing two components whose frequency ratio is high. The diagram (*c*) is a waveform which is seen to contain the 1st and 3rd harmonics, with an indication of some higher harmonic. If the process described in formula (2.8) is performed on the original variation (3.1), the result will disclose those harmonics whose reference numbers are multiples of 3. The numerical work is shown in Table II. Columns (*a*), (*b*) and (*c*)

TABLE II

(*a*)	(*b*).	(*c*).	(*d*) = (*a*) + (*b*) + (*c*)	$e = \frac{1}{3}(d)$	(*f*)
39·4	17·7	15·6	72·7	24·2	4·2
29·4	27·6	21·0	78·0	26·0	6·0
26·5	18·3	27·9	72·7	24·2	4·2
26·5	14·0	19·5	60·0	20·0	0
17·1	13·9	16·3	47·3	15·8	−4·2
15·0	10·9	16·1	42·0	14·0	−6·0
12·6	19·5	15·2	47·3	15·8	−4·2
6·3	23·7	30·0	60·0	20·0	0
			72·7	24·2	4·2
			etc.	etc.	etc.

list the values in the three ranges 0 to 105°, 120° to 225° and 240° to 345°. The sums of corresponding figures in the three columns are given under (*d*), and these values are divided by 3 under (*e*). These last values are plotted in Fig. 1*d*, and it can be seen at once that the diagram represents the function $20 + 6 \sin 3(x + 15°)$. Column (*f*) in Table II lists the values obtained by subtracting the constant term 20 from column (*e*); the validity of the conclusion that the series of the 3rd, 6th, 9th, etc., harmonics consists solely of the 3rd harmonic $6 \sin 3(x + 15°)$ can be checked from the fact that at $x = 0$, this harmonic has the value $6 \times 0{\cdot}707 = 4{\cdot}2$, correct to one decimal place, and this value agrees with the figure in column (*f*).

Column (*f*) of Table II therefore represents the 3rd harmonic; the result of subtracting it from column (*f*) of Table I would be to

indicate the sum of the remaining odd harmonics. Table III gives the numerical working for the first half-cycle, the figures for the remaining half-cycle being derived from the fact that the result must be alternant, since it includes only odd harmonics (see Chapter II, p. 58), so that the value for $x = 270°$ will be equal in magnitude and opposite in sign to that for $x = 90°$, and so on.

TABLE III

(*a*) = I(*f*)	(*b*) = II(*f*)	(*c*) = (*a*) (*b*)
12·8	4·2	8·6
9·3	6·0	3·3
3·5	4·2	−0·7
1·4	0	1·4
0·8	−4·2	5·0
− 3·0	−6·0	3·0
− 8·7	−4·2	−4·5
− 6·6	0	−6·6
0·7	4·2	−3·5
5·8	6·0	−0·2
1·6	4·2	−2·6
− 8·0	0	−8·0
−12·8	−4·2	−8·6
etc.	etc.	etc.

The result is plotted in Fig. 1*e*, from which it can at once be seen that the remaining odd harmonics are the 1st and the 5th; the additional information that may be required concerning these harmonics (amplitudes and phase-angles) may be obtained by envelope analysis.

More complete information can be obtained numerically by separating the sine and cosine components of the waves in Fig. 1, *b* and *c*. Table IV gives the separation for the wave in Fig. 1*b*, by application of the process represented by formulæ (2.13) to the figures of column (*d*) of Table I, after first subtracting from each of these figures the constant term 20. It will be noted that as the variation now being considered contains only the even harmonics, it is sufficient to include only a half-cycle in Table IV.

Now, column (*d*) of Table IV represents the function

$$y = a_2 \cos 2x + a_8 \cos 8x.$$

At $x = 0$, $y = a_2 + a_8 = 6{\cdot}7$; and at $x = 90°$, $y = -a_2 + a_8 = 0{\cdot}3$. Hence $a_2 = 3{\cdot}2$, $a_8 = 3{\cdot}5$.

TABLE IV

x°	[I(d) 20]		(c) = (a) + (b)	(d) = ½(c)	(e) = (a) − (b)	(f) = ½(e)
	(a)	(b)				
0	6·7	6·7	13·4	6·7	0	0
15	0·2	2·0	2·2	1·1	−1·8	−0·9
30	3·0	−3·2	−0·2	0·1	6·2	3·1
45	5·1	1·9	7·0	3·5	3·2	1·6
60	−3·6	−3·0	−6·6	−3·3	−0·6	−0·3
75	−2·0	−7·1	−9·1	−4·6	5·1	2·6
90	0·3	0·3	0·6	0·3	0	0
				−4·6		−2·6
				etc.		etc.

Similarly, column (f) represents the function

$$y = b_2 \sin 2x + b_8 \sin 8x\,;$$

at $x = 45^\circ$, $y = b_2 = 1{\cdot}6$, and at $x = 75^\circ$,

$$y = b_2 \sin 150^\circ + b_8 \sin 600^\circ = 0{\cdot}8 - 0{\cdot}866\, b_8 = 2{\cdot}6.$$

Hence $b_8 = -\,1{\cdot}8/0{\cdot}866 = -\,2{\cdot}1$.

The amplitudes of the two harmonics are therefore given by

$$A_2 = \sqrt{a_2^2 + b_2^2} = 3{\cdot}6,$$
$$A_8 = \sqrt{a_8^2 + b_8^2} = 4{\cdot}1,$$

and the phase-angles are determined by

$$\tan \phi_2 = a_2/b_2 = 2{\cdot}0, \quad \phi_2 = 63^\circ \text{ approx.},$$
$$\tan \phi_8 = a_8/b_8 = -\,1{\cdot}67, \ \phi_8 = 139^\circ \text{ approx.}$$

(See note in Chapter I, p. 15, on the determination of phase-angles by this means.)

The amplitudes and phase-angles of the 1st and 5th harmonics can be found in this manner, and the calculation is left as an exercise for the reader.

The determination of the sine or cosine component of a harmonic wave from a wave which is known to consist of the sine or cosine components of two harmonics can always be performed in this manner. The procedure suffers, however, from a disadvantage, in so far as the particular values taken for the determination of the amplitudes may be subject to errors which render the result quite inaccurate. The disadvantage could be overcome by determining the amplitudes from numerous different pairs of values, and averaging the results, on the principle that any errors inherent in the data would tend to average out; this process is really the basis of the purely numerical

method of analysis, described in Chapter VII, and if the analyst is prepared to expend the time required by such a procedure in this special application he might as well use the purely numerical routine method, which will give the best possible results obtainable from the data.

4. Frequency ratios 2 : 1, 3 : 1, 3 : 2, etc.

The method of superposition is particularly useful in the analysis of waveforms containing two or three components whose frequency ratios are in the series 3 : 2 : 1, or some similar simple series ; these waves cannot usually be analysed successfully by the envelope method, and hardly justify the employment of the routine purely numerical method to be described hereafter. The general procedure consists of analysing the variation into two simpler variations, containing, respectively, the even and the odd harmonics, by graphical or numerical superposition of two half-cycles, and separating out the harmonics whose reference numbers are multiples of 3 by graphical or numerical superposition of three equal sub-cycles.

The numerical process of superposition has been adequately illustrated in the preceding section, and the graphical method will now be described.

Fig. 2 shows the application of the first part of the procedure to a particular wave, illustrated at (*a*). In the absence of any indication as to the zero-level or mean line, an arbitrary datum is drawn in as at AA′; the direction of this line is determined by the fact that it must be parallel to any line joining corresponding points in successive cycles. The cycle AC is divided into two equal parts AB and BC, the two parts being superimposed as at (*b*). The full line in (*b*) shows the mean of the variations represented by the broken lines AB and BC, and is in fact the second harmonic added to the constant term.

This superposition is best performed with the aid of tracing paper. The second half-cycle BC is traced off and placed in the correct position over the first half-cycle AB ; the derivation of the mean (full) line can then be achieved by either of two methods :

(i) With proportional dividers, corresponding ordinates to the two half-cycles are added with the dividers set at the 2 : 1 ratio, and half the sum is marked off from the datum-line.

(ii) A piece of opaque paper with a straight edge is moved over the traces, with the edge perpendicular to the datum-line. At each position a point is marked on the tracing

paper so that it bisects the segment of the edge of the movable paper cut off by the two half-cycles, the point of bi-section being estimated by eye. This method may appear at first sight to be crude, but very accurate results can be obtained with a little practice.

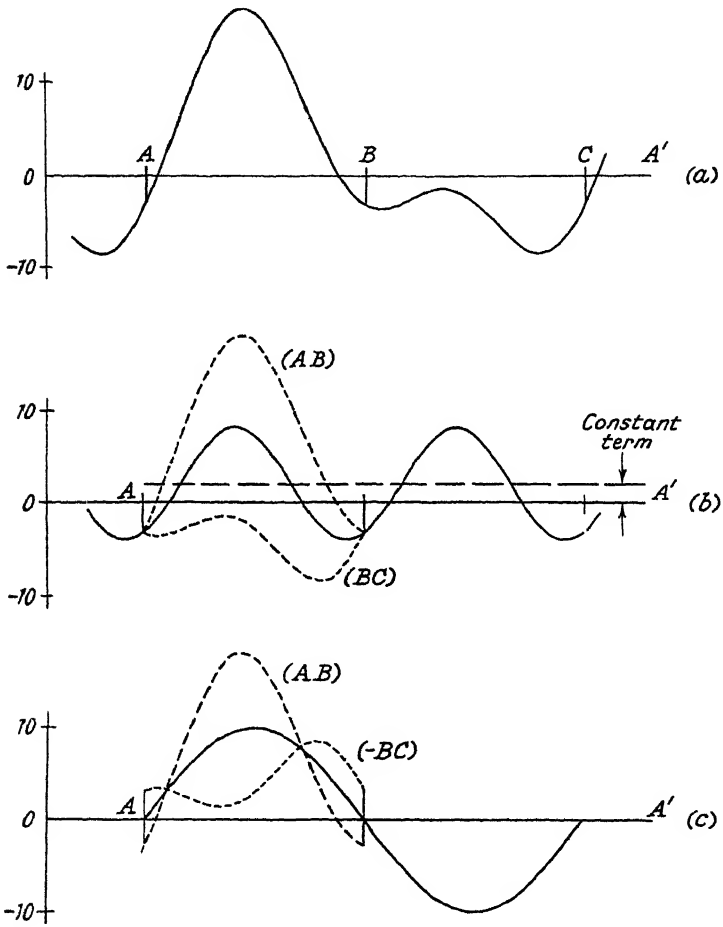


FIG. 2.—Graphical version of superposition method. (a) Waveform to be analysed; (b) even harmonics, found by graphical addition of two half-cycles; (c) odd harmonics, found by graphical subtraction of two half-cycles.

At (c) in the diagram the part AB of the cycle has been redrawn, and the part B has been superimposed after inversion (i.e. after the sign of each ordinate has been changed. The full line is again

the mean of the other two, and represents in this case the sum of the odd harmonics—the first harmonic only. The original tracing of the part BC can be utilised for this purpose, the paper being simply turned over so as to bring about the desired inversion.

As a result of this procedure, the waveform has been completely analysed; the function represented by it is

$$y = 2 + 10 \sin x + 6 \sin 2(x - 30°).$$

Again, in Fig. 3 the wave is seen to have only odd harmonics,

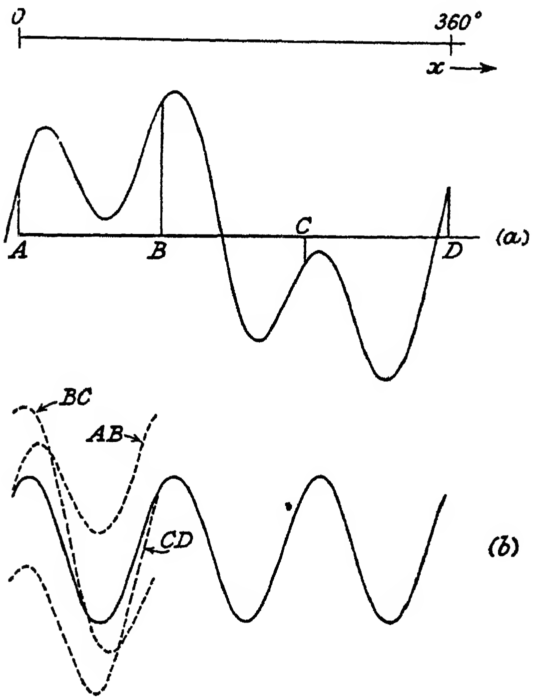


FIG. 3.—Graphical separation of third harmonic.

since it is alternant (Chapter II, p. 58). The line AD is the mean line, and in Fig. 3*b* the three equal sub-cycles AB, BC and CD are redrawn; the full line represents the average of the variations represented by the three broken lines, and is seen to be the third harmonic. The averaging process in this case is best performed by means of proportional dividers, the three ordinates being summed with the dividers set to 3 : 1; but it is quite practicable to use the eye-estimation method, a point being marked on each ordinate at the mean position between two of the values, and the

final point being selected so as to divide the distance between this point and the third value in the ratio 1 : 2.

By subtracting the third harmonic from the original wave the remaining harmonics are indicated, and this is left as an exercise for the reader.

In Fig. 4*a* is shown a fairly simple waveform which is not symmetrical, skew-symmetrical or alternant. Half-cycles of the even and odd harmonics are derived, at (*b*) and (*c*), by the methods already described. The full trace at (*b*) is the 2nd harmonic, and the full trace at (*c*) is the same as the first half of the waveform

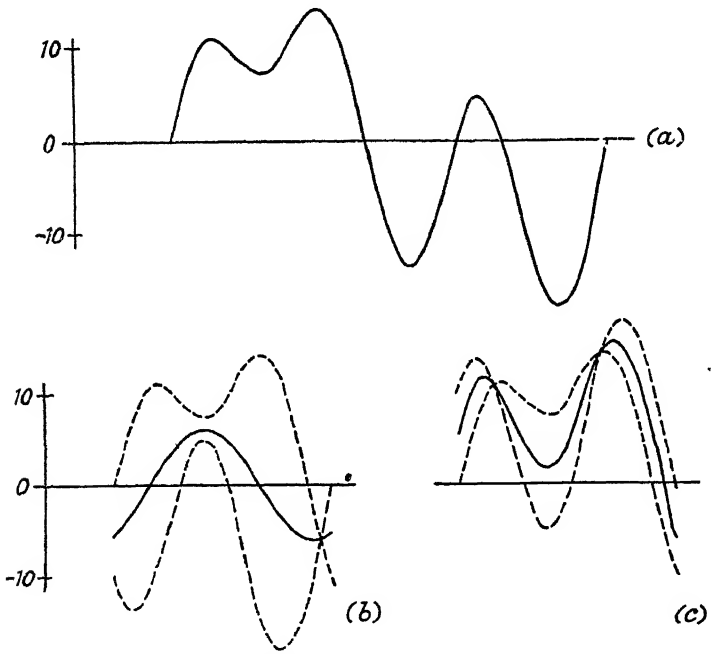


FIG. 4.—Further example of graphical analysis.

at Fig. 3(*a*); thus the waveform at (*a*) is a combination of a 2nd harmonic (*b*) and the waveform Fig. 3*a* which has already been analysed.

The constant term can in any case be determined by the fact that all harmonic components have a zero average value over a complete cycle, so that the constant term A_0 in the series (2.1) is the average value of the function $f(x)$ over the complete cycle. In the case of alternant or skew-symmetrical waveforms the mean line can at once be determined by the fact of the alternance or skew-symmetry.

5. Extension of method.

By a further extension of the method of superposition it is possible to determine most of the components of a waveform which contains harmonics up to, say, the twelfth.

If the variation is given in the form of numerical values at regular intervals over the cycle, the constant term A_0 is found by averaging, as explained above. If, however, the waveform is recorded, the determination of the constant term is deferred to a later stage.

For convenience a new notation will be employed in the description of the method. The result of splitting a cycle of a trace (a) into n sub-cycles, adding them together and dividing by n, will be denoted by $I_n(a)$. Thus in Fig. 3 the trace (b) is derived from (a) by the operation $(b) = I_3(a)$, and in Fig. 4 the trace (b) is derived from (a) by the operation $(b) = I_2(a)$. The process of subtracting the second half-cycle from the first half-cycle and halving the result will be denoted by I_2'; thus in Fig. 4, $(c) = I_2'(a)$.

Referring to the various harmonics by their reference numbers, and using zero to denote the constant term, the original waveform (y) containing a constant term and all the first twelve harmonics is represented as follows:

$$(y) = (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12). \quad . \quad (5.1)$$

Then
$$(a) = I_2(y) = (0, 2, 4, 6, 8, 10, 12)$$
$$(b) = I_2'(y) = (1, 3, 5, 7, 9, 11).$$
$$(c) = I_3(y) = (0, 3, 6, 9, 12).$$
$$(d) = I_4(y) \text{ or } I_2(a) = (0, 4, 8, 12).$$

In these expressions the numbers inside the brackets on the right-hand side represent the harmonic contents of the functions named on the left-hand side.

Proceeding further,

$$(e) = I_2'(c) \text{ or } I_3(b) = (3, 9). \quad . \quad . \quad . \quad (5.2)$$

[It should be noted that in all cases one cycle of the variation which is the subject of the operation is treated. Thus, for example, (c) has three cycles in one cycle of (y), and in deriving (e) the parts of the wave which are subtracted are halves of these cycles, i.e. sixths of the original cycle. On the other hand, (b) has the same wavelength as (y); but the final result is the same. It is assumed here that all the first twelve harmonics are present: the procedure to be adopted in the simpler cases where some of the harmonics are absent will be evident: the difficulty lies rather in representing

the process concisely in symbolic form.] The sum of the 3rd and 9th harmonics has thus been determined ; if either of these is zero the other is given at once by (5.2), and if both are present they can be separated by the process

$$(f) = \mathrm{I}_3(e) = (9), \qquad (5.3)$$

and
$$(g) = (e) - (f) = (3). \qquad (5.4)$$

Now,
$$(h) = (c) - (e) = (0,\ 6,\ 12). \qquad (5.5)$$

If either the 6th or the 12th harmonic is absent the other is given immediately, together with the constant term, by (5.5). Otherwise, the process I_2 on (h) will yield the required information :

$$(i) = \mathrm{I}_2(h) = (0,\ 12), \qquad (5.6)$$

and
$$(j) = (h) - (i) = (6). \qquad (5.7)$$

These last equations enable both the harmonics and the constant term to be determined.

Further,
$$(k) = (d) - (i) = (0,\ 4,\ 8). \qquad (5.8)$$

If either the 4th or the 8th harmonic is absent, the equation (5.8) enables the other to be determined at once ; otherwise the process I_2 on (k) yields the 8th harmonic :

$$(l) = \mathrm{I}_2(k) = (0,\ 8), \qquad (5.9)$$

and
$$(m) = (k) - (l) = (4). \qquad (5.10)$$

The following harmonics have so far been determined : 0, 3, 4, 6, 8, 9, 12.

Let
$$(n) = (h) + (k) - (0) = (0,\ 4,\ 6,\ 8,\ 12),$$

and
$$(p) = (a) - (n) = (2,\ 10). \qquad (5.11)$$

If either the 2nd or the 10th harmonic is absent the other is given at once by (5.11) ; otherwise the function (5.11) can be plotted and analysed into its components by inspection (envelope analysis), or by separating the sine and cosine components as described on page 127 in connection with Fig. 1. . . . . . . (5.12)

Finally,
$$(q) = (b) - (e) = (1,\ 5,\ 7,\ 11). \qquad (5.13)$$

In the event of all four of these harmonics being present, the processes I_5 and I_7 yield, respectively,

$$(r) = \mathrm{I}_5(q) = (5), \qquad (5.14)$$

$$(s) = \mathrm{I}_7(q) = (7), \qquad (5.15)$$

and
$$(q) - (r) - (s) = (1,\ 11). \qquad (5.16)$$

The last function can be analysed into its components by the inspection method. It is more probable, however, that only two of the four harmonics will be present in (q). If these two are (1, 5), (1, 7) or (1, 11) the wave can be analysed by the envelope method as a wave of two components whose frequency ratio is high; if (5, 7), (5, 11) or (7, 11) it would be better to perform the operation I_5 or I_7.

6. Theorem.

The theorem quoted on page 123, above, is here proved. In the course of the proof use is made of the elementary properties of complex numbers, for a short informal treatment of which see reference 2 in the Bibliography at the end of the book.

Let $S = \sin\theta + \sin(\theta - 2\pi/n) + \sin(\theta - 4\pi/n) + \ldots$

$$\ldots + \sin[\theta - (n-1)(2\pi/n)],$$

and $C = \cos\theta + \cos(\theta - 2\pi/n) + \cos(\theta - 4\pi/n) + \ldots$

$$\ldots + \cos[\theta - (n-1)(2\pi/n)].$$

Then if

$$i^2 = -1, \quad C + iS = z + zr + zr^2 + \ldots\ldots + zr^{n-1},$$

where $\quad z = e^{i\theta} \quad$ and $\quad r = e^{i2\pi/n}$.

Hence
$$C + iS = \frac{z(1-r^n)}{1-r},$$

$$= \frac{z}{1-r}[1 - \cos 2\pi + i\,.\,\sin 2\pi] = 0.$$

Hence $C = S = 0$; in particular,

$\sin\theta + \sin(\theta - 2\pi/n) + \sin(\theta - 4\pi/n) + \ldots$

$$\ldots + \sin[\theta - (n-1)(2\pi/n)] = 0.$$

CHAPTER VI

FOURIER SERIES: MATHEMATICAL ANALYSIS

1. Introductory.

Fourier may be said to have been the real founder of the mathematical study of periodic functions. Although for nearly fifty years before he presented his first paper to the Paris Academy in 1807 there had been a protracted controversy between D. Bernoulli, Euler, Lagrange, d'Alembert and others on the subject of the solution to the problem of vibrating strings, in the course of which Lagrange almost stumbled upon the discovery which later made Fourier famous,* yet the vital discovery was made by Fourier in the course of an entirely different investigation. His interest in the mathematical theory of the conduction of heat was responsible for the formulation of the celebrated theorem which has proved to be a powerful tool in the solution of many widely different types of problems.

It may be noted that the controversy which preceded and followed Fourier's contribution to the subject concerned the possibility of expanding *arbitrary* functions in the form of trigonometrical series. It was agreed that an analytic periodic function, i.e. one which could be represented by a single mathematical expression over the entire period, can be so expanded; the argument concerned what Eagle terms "artificial functions," which are represented by different analytic functions over different parts of the cycle. For further details of the history of the subject, the reader is referred to the works of Eagle and Carslaw, references 1 and 2 in the Bibliography.

Perhaps the most important application of Fourier's Theorem is to vibration theory, and the fact that Lagrange was engaged on the problem of vibrating strings when he so nearly anticipated Fourier's discovery is doubly interesting for that reason. Since the beginning of the nineteenth century the theorem has been applied in practically every branch of mathematical physics, and no textbook on mathematical methods in engineering problems is considered complete without a lengthy treatment of the theorem and its consequences.

* Lagrange gave a formula in which it is only necessary to make two slight modifications—one a simplification—in order to arrive substantially at Fourier's Theorem.

The theorem in its simplest form states that any periodic variation fulfilling certain conditions regarding continuity can be considered as the sum of a number of sinusoidal variations whose periods exhibit a simple relationship: furthermore, for any given function or variation the equivalent series of sinusoidal variations is unique. The theorem does not state that the original variation must *necessarily* be regarded as such a series; in fact, modern mathematics displays a large class of functions (the *orthogonal functions*) of which the sine function is the simplest, and any of which can be used as the basis of a series representation of arbitrary functions. However, the fact that the free undamped motion of a simple vibrating system having one degree of freedom can be represented as a sine function of time, together with the fact that the alternating electric current delivered from a dynamo is likewise sinusoidal, concentrates the attention of practical investigators on the representation of periodic variations by series of sinusoidal variations—i.e. by Fourier series. This matter is discussed more fully in Appendix I, page 244.

The mathematical treatment of the subject, which is given in the present chapter, is of more than academic interest. In very many practical cases the methods here described of analysing a complex variation into its component sine-waves are the only methods applicable; furthermore, a thorough understanding of these methods is essential to the intelligent use of any mechanical aids to analysis, such as are described in Chapter VIII. An elementary knowledge of the integral calculus is required, but beyond that the engineer or research technician may safely leave the theoretical niceties to the professional mathematicians; as Heaviside has remarked, one can tell the time by one's watch without having a detailed knowledge of the mechanism.

In this connection it is interesting to note that Fourier himself, who was a scientist rather than a mathematician, never *proved* the theorem to the satisfaction of contemporary mathematicians, who flatly denied the possibility of expanding an arbitrary function in the form of a Fourier series.

Certain considerations of the continuity of functions are required, and these are discussed briefly at the beginning of the chapter. Fourier's Theorem is then stated, and the method of determining the coefficients in the series by means of the Cauchy integrals explained, the truth of the theorem being assumed. (An interesting and ingenious "demonstration" of the theorem, given by the late Professor Donkin, is quoted in Appendix IV, p. 253.) To illustrate the procedure, several simple waveforms of importance in electrical

theory are treated at length. In Section 6 the concept of *periodic existence functions* is developed briefly; certain aspects of this concept are thought to be novel, and apart from the simplification which their use produces in the analysis of certain types of waveform it may well be that these functions can be applied in other fields of investigation: for these reasons it has been considered advisable to include this short account.

The convergence of Fourier series is discussed informally in Section 7, and an account is given of Gibbs' phenomenon in a particular example. The chapter concludes with a reference to other forms of Fourier series which have been found of great use in engineering problems.

Limitations of space prevent a full treatment of the theory of Fourier series from being given, and neither is this necessary for the purposes of the present work. The reader who desires a more extensive or more rigorous discussion of the general properties of these series is referred to the works of Carslaw, and Whittaker and Watson, references 3 and 4 in the Bibliography.

2. Functions: limits and continuity.

The functions which are the subject of harmonic analysis are single-valued and fulfil certain conditions of continuity; these concepts are examined in this section.

A functional equation represents a relationship between two or more variable quantities. Thus the equation

$$y = f(x), \qquad . \quad . \quad . \quad . \quad (2.1)$$

which is read "y is a function 'f' of x," means that as x is given different values so definite values of y are indicated according to the functional relationship "f." If $f(x) = x^2$, then as x is given a range of values y takes a definite range of values according to the equation

$$y = x^2. \qquad . \quad . \quad . \quad . \quad (2.2)$$

In the functional form (2.1) x is said to be the *independent variable*, *basic variable*, or *argument* of the function y; this means that certain values are assigned to x, and the corresponding values of y are calculated from (2.1). As the values of y depend upon the values of x, y is said to be the *dependent variable*. In many cases it is possible to express the equation connecting x and y (of which 2.2 is an example) in such a way as to define x in terms of y, say

$$x = \mathrm{F}(y), \qquad . \quad . \quad . \quad . \quad (2.3)$$

so that the dependent and independent variables are interchanged. The functional form F is then said to be the *inverse* to the functional form f; and this is often expressed symbolically in the following manner:

$$F = f^{-1}. \qquad \qquad (2.4)$$

Of this last equation the trigonometric inverse functions are well-known examples; if $y = \sin x$, the equation giving x in terms of y is put in the form

$$x = \sin^{-1} y.$$

The function y is said to be a *single-valued* or a *multi-valued* function of x according as one, or more than one, value of y is determined by the equation $y = f(x)$. Thus the equation (2.2) defines y as a single-valued function of x, since for any value assigned to x there is one and only one corresponding value determined for y. On the other hand, if (2.2) is rewritten as

$$x = y^{\frac{1}{2}},$$

x is now defined as a double-valued function of y, since for every value assigned to y there are two values indicated for x, since

$$(-x)^2 = x^2.$$

The functions which express the relation between the co-ordinates of points on waveforms are single-valued, in that the wave has only one point corresponding to any particular value of the basic variable.

The concept of *continuity*, applied to functions, is capable of visual appreciation when the graph depicting the function is studied. Thus the curve in Fig. 1*a* is "obviously" continuous over the range plotted, while that in Fig. 1*b* is "obviously" discontinuous at the value $x = X$. To obtain an analytic definition of continuity it is necessary first to consider *limits*.

The function $f(x)$ is said to "tend to the limit l when x tends to X from above" if, no matter how small δ is chosen, a quantity ϵ can be found such that

$$|f(x) - l| < \delta$$

whenever $$X < x \leqslant X + \epsilon.$$

(In this formula the sign | | denotes "numerical value of.") What constitutes a suitable value of ϵ depends on the preliminary choice of δ, and to express this fact the quantity is usefully written in functional form as $\epsilon(\delta)$. The fact that l is the limiting value of

$f(x)$ as x tends to the value X from above is stated symbolically thus :

$$\underset{x \to X+0}{\text{Lt.}} f(x) = l, \qquad . \quad . \quad . \quad . \qquad (2.5)$$

in which the notation $x \to X + 0$ denotes a restriction of attention to values of x greater than X. The complete statement therefore runs :

"$\underset{x \to X+0}{\text{Lt.}} f(x) = l$ if, given a value of δ no matter how small, a quantity $\epsilon(\delta)$ can be determined such that

$$|f(x) - l| < \delta \text{ when } X < x \leqslant X + \epsilon(\delta)." \qquad (2.6)$$

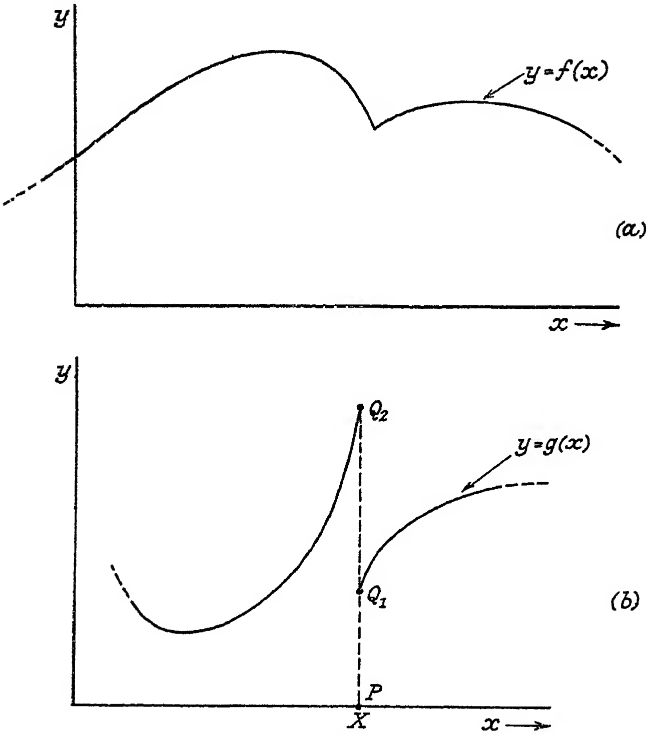


FIG. 1.—Continuity and finite discontinuity. (*a*) A continuous function; (*b*) a function with a finite discontinuity Q_1Q_2.

Similarly, the condition for the existence of a limiting value l of $f(x)$ as x tends to the value X from below is expressed thus :

"$\underset{x \to X-0}{\text{Lt.}} f(x) = l$ if, given a value of δ no matter how small, a quantity $\epsilon(\delta)$ can be found such that

$$|f(x) - l| < \delta \text{ when } X - \epsilon(\delta) \leqslant x < X." \quad . \quad (2.7)$$

Frequently for brevity the two expressions for the limits are written in the form:

$$f(X+0) = \underset{x \to X+0}{\text{Lt.}} f(x),$$
$$f(X-0) = \underset{x \to X-0}{\text{Lt.}} f(x),$$

and if $f(X+0) = f(X-0)$, the simple form $f(X)$ suffices to indicate the value of $f(x)$ when $x = X$.

Two points in connection with these definitions should be observed. First, the use of the modulus sign (| |) shows that it is immaterial whether $f(x)$ is an increasing or a decreasing function of x in the neighbourhood of $x = X$, since the numerical value without sign of the difference $f(x) - l$ is to be taken. Secondly, in the development of an analytic condition for continuity it is important to note that both the limits (2.6) and (2.7) may exist (and may be equal) and yet the function be undefined for the value $x = X$. Thus if

$$f(x) = x \,.\, \sin \frac{1}{x},$$

then, since the sine function never takes any value greater numerically than unity,

$$\underset{x \to 0}{\text{Lt.}} f(x) = 0,$$

but $f(0)$ is undefined, since no meaning has been assigned to the expression $\sin \infty$. The notation $x \to X$ is commonly employed if

$$f(X+0) = f(X-0).$$

The "obvious" discontinuity of the function represented in Fig. 1*b* at $x = X$ is now seen to be due to the inequality of the two limits $g(X+0)$ and $g(X-0)$, for

$$g(X+0) = PQ_1,$$
and
$$g(X-0) = PQ_2$$

and Q_1, Q_2 are not coincident. With the condition, evidently necessary, that for $f(x)$ to be continuous at the value $x = X$ it must be defined for this value, so that $f(X)$ has a determinate value, the definition of continuity at a point is obtained:

> "The function $f(x)$ is said to be continuous at the value $x = X$ if it tends to a limit as x tends to X from either side, and each of these limits equals $f(X)$."

From the definitions (2.6), (2.7) of the limits an alternative form of statement is obtained:

"$f(x)$ is continuous at $x = X$ if, given δ no matter how small, a quantity $\epsilon(\delta)$ can be found such that

$$|f(x) - f(X)| < \delta \text{ if } 0 < |x - X| \leqslant \epsilon(\delta)." \quad . \quad (2.8)$$

To demonstrate the discontinuity of the function depicted in Fig. 1*b*, according to this definition, suppose that $g(x)$ is defined

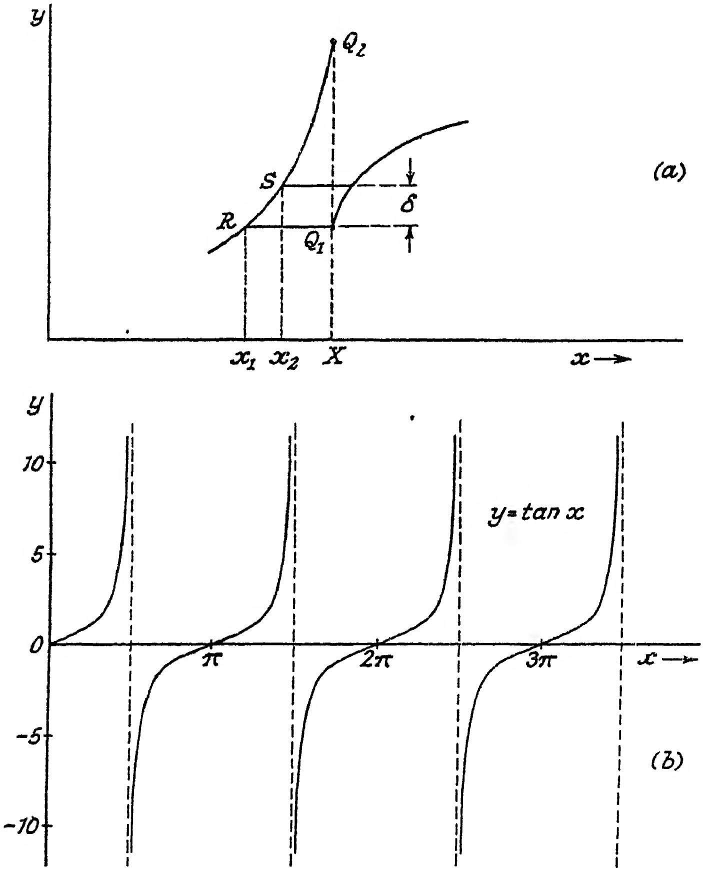


FIG. 2.—Discontinuities: (*a*) finite; (*b*) infinite.

to be equal to PQ_1 at $x = X$. Then if δ be chosen less than Q_1Q_2, it is not possible to find the required quantity $\epsilon(\delta)$; for although any value of x in the range from x_1 to x_2 (Fig. 2*a*) gives a point in RS such that $|g(x) - g(X)| < \delta$, any value of x in the range from x_2 to X gives a point in SQ_2 such that $|g(x) - g(X)| > \delta$, and the condition "if $0 \leqslant |x - X| \leqslant \epsilon(\delta)$" is not fulfilled.

It is now possible to define *continuity over a range*. The function

$f(x)$ is said to be continuous over the range $a < x < b$ if it is continuous for all values of x in that range. From the preceding remarks it can be seen that the function may be continuous over a range without being defined for the end-points $x = a$ and $x = b$; and that an essential condition for continuity over a range is that the function be defined for every value of x in the range.

The function depicted in Figs. 1*b* and 2*a* is *finitely* discontinuous at the value $x = X$, since the difference between the limits $g(X + 0)$ and $g(X - 0)$ is finite; in such circumstances X is said to be a value of *ordinary discontinuity.* On the other hand, the function $\tan x$ is *infinitely* discontinuous at the values $x = \pi/2$, $3\pi/2$, etc., since for example $\tan(\pi/2 - 0) = \infty$ while $\tan(\pi/2 + 0) = -\infty$, as shown in Fig. 2*b*. The distinction between ordinary and infinite discontinuity must be preserved; functions involving infinite discontinuities require special treatment in Fourier Analysis.

3. Fourier's Theorem.

"If a function $f(x)$ of x has a finite number of points of ordinary discontinuity and a finite number of maxima and minima in the interval $0 \leqslant x \leqslant 2\pi$, and if the function be arbitrarily defined within this interval and defined by the relation

$$f(x + 2\pi) = f(x), \qquad (3.1)$$

for values of x outside this interval, then

$$f(x) = \tfrac{1}{2}a_0 + \sum_{k=1}^{k=\infty}(a_k \cos kx + b_k \sin kx), \qquad (3.2)$$

where

$$\left.\begin{aligned} a_k &= \frac{1}{\pi}\int_0^{2\pi} f(x) \cos kx \,.\, dx \\ b_k &= \frac{1}{\pi}\int_0^{2\pi} f(x) \sin kx \,.\, dx \; (k = 0, 1, 2, \text{etc.}) \end{aligned}\right\} \qquad (3.3)$$

and at every point $x = X$ in the interval the series (3.2) converges to the value

$$f(X) = \tfrac{1}{2}[f(X + 0) + f(X - 0)]." \qquad (3.4)$$

This is a statement of Fourier's Theorem, named after the celebrated French scientist. Fourier's first paper on the subject of trigonometric series was read to the French Academy in 1807. He discovered the theorem in his researches into the conduction of heat. The series (3.2) is known as the *Fourier series* or *Fourier expansion* of the function $f(x)$; the coefficients a_k and b_k are termed the *Fourier coefficients*, and the process of determining these coefficients is called *Fourier analysis* or *harmonic analysis.*

The conditions stated are unnecessarily stringent. Dirichlet

established various sets of conditions during the years 1829-1837, but it is probable that the necessary and sufficient set of conditions has not yet been formulated. Infinite discontinuities can be dealt with, but the case is seldom met in practical work and the accuracy of the series for values of x in the neighbourhood of such discontinuities is not great unless a very large number of terms is taken.

The equation (3.1) states that $f(x)$ is a periodic function of x, with the period 2π. Periodic functions with periods other than 2π can easily be converted into this form, as described in Section 2 of Chapter II. The equation (3.4) states that for points of continuity the series converges to the value of the function, while for points of finite discontinuity the series converges to the mean of the two limits $f(X - 0)$ and $f(X - 0)$. In the present section the truth of equation (3.4) is assumed, and the quantity $f(X)$ given by substituting $x = X$ in the series (3.2) will be taken as agreeing with the arbitrary function $f(x)$ at every point $x = X$. General considerations of the convergence of Fourier series are set out in Section 7 below.

The equations (3.3) are due to Cauchy. It is to be noted that, in order to enable the constant term in the series to be calculated from the same formula as the coefficients of the cosine terms, the symbol A_0 of equation (4.2) of Chapter II has been replaced by $\frac{1}{2}a_0$ In the following proof of the *Cauchy integrals* (3.3) it is assumed that it is legitimate to integrate a series such as (3.2) term by term. The consideration of this point is too involved to be included here, but reference may be made to any comprehensive textbook on mathematical analysis or the theory of functions (see references A, C in the Bibliography at the end of the book).

Determination of coefficients (*Cauchy integrals*). In the proof of the equations (3.3) the following results are required :

$$\left.\begin{array}{l}
\left.\begin{array}{ll}
(a) & \int_0^{2\pi} \sin mx \,.\, \sin nx\, dx = 0 \\
(b) & \int_0^{2\pi} \sin mx \,.\, \cos nx\, dx = 0 \\
(c) & \int_0^{2\pi} \cos mx \,.\, \cos nx\, dx = 0
\end{array}\right\} \text{if } m \text{ and } n \text{ are unequal integers.} \\
\left.\begin{array}{ll}
(d) & \int_0^{2\pi} \sin kx \,.\, \cos kx\, dx = 0 \\
(e) & \int_0^{2\pi} \sin kx\, dx = 0 \\
(f) & \int_0^{2\pi} \cos kx\, dx = 0 \\
(g) & \int_0^{2\pi} \sin^2 kx\, dx = \pi \\
(h) & \int_0^{2\pi} \cos^2 kx\, dx = \pi
\end{array}\right\} \text{if } k \text{ is an integer}
\end{array}\right\} \qquad (3.5)$$

These results are easily obtained. By ordinary trigonometrical substitution,

$$\sin mx \,.\, \sin nx = \tfrac{1}{2}\cos(m-n)x - \tfrac{1}{2}\cos(m+n)x,$$

so that if m and n are unequal,

$$\int_0^{2\pi} \sin mx \,.\, \sin nx\, dx = \tfrac{1}{2}\int_0^{2\pi} \cos(m-n)x\, dx - \tfrac{1}{2}\int_0^{2\pi} \cos(m+n)x\, dx$$

$$= \frac{1}{2(m-n)}\Big[\sin(m-n)x\Big]_0^{2\pi} - \frac{1}{2(m+n)}\Big[\sin(m+n)x\Big]_0^{2\pi}.$$

Now as both m and n are integers, the quantities $\sin(m-n)x$ and $\sin(m+n)x$ both have zero value at the upper and lower values of x ($= 2\pi$ and 0). Hence both terms in the last equation are zero, and thus equation (3.5a) is established.

If, however, $m = n = k$, say, then as

$$\cos 2kx = 1 - 2\sin^2 kx,$$

$$\int_0^{2\pi} \sin^2 kx\, dx = \tfrac{1}{2}\int_0^{2\pi} (1)\, dx - \tfrac{1}{2}\int_0^{2\pi} \cos 2kx\, dx$$

$$= \tfrac{1}{2}[x]_0^{2\pi} = \pi,$$

since the second integral is zero. Thus equation (3.5g) is established. This result may be obtained directly from the proof of (3.5a) given immediately above, by utilising L'Hospital's rule. As this rule is sometimes useful in the determination of Fourier coefficients, it is now given and proved.

"If $f(t)$ and $F(t)$ are both zero when $t = T$, the limit of the ratio $f(t)/F(t)$ as t tends to T is equal to the limit of the ratio $f'(t)/F'(t)$ as t tends to T, where dashes denote differentiation with respect to t."

Let $t = T + \delta T$, then expanding by Taylor's Theorem,

$$\underset{t \to T}{\text{Lt.}} \frac{f(t)}{F(t)} = \underset{\delta t \to 0}{\text{Lt.}} \frac{f(T+\delta T)}{F(T+\delta T)}$$

$$= \underset{\delta T \to 0}{\text{Lt.}} \frac{f(T) + \delta T \,.\, f'(T) + \frac{1}{2}(\delta T)^2 f''(T) + \text{etc.}}{F(T) + \delta T \,.\, F'(T) + \frac{1}{2}(\delta T)^2 F''(T) + \text{etc.}}$$

Now $f(T) = F(T) = 0$, hence

$$\underset{t \to T}{\text{Lt.}} \frac{f(t)}{F(t)} = \underset{\delta T \to 0}{\text{Lt.}} \frac{f'(T) + \frac{1}{2}\delta T \,.\, f''(T) + \text{etc.}}{F'(T) + \frac{1}{2}\delta T \,.\, F''(T) + \text{etc.}}$$

$$= \underset{\delta T \to 0}{\text{Lt.}} \frac{f'(T)}{F'(T)} = \underset{t \to T}{\text{Lt.}} \frac{f'(t)}{F'(t)}. \qquad (3.6)$$

Applying this rule to the last line in the proof of equation (3.5*a*) above, the first term on the right-hand side becomes, when $m \to n$,

$$\left[\underset{m\to n}{\mathrm{Lt.}}\frac{\sin (m-n)x}{2(m-n)}\right]_0^{2\pi} = \left[\underset{m\to n}{\mathrm{Lt.}}\frac{x \cos (m-n)x}{2}\right]_0^{2\pi},$$

since $$\frac{d}{dm}[\sin (m-n)x] = x \cos (m-n)x,$$

and $$\frac{d}{dm}[2(m-n)] = 2.$$

The value of this last limit is clearly $\frac{1}{2}x$, and

$$\left[\tfrac{1}{2}x\right]_0^{2\pi} = \pi.$$

The remaining equations of (3.5) are proved in a similar manner. For equations (*b*) and (*c*) the trigonometrical results are used :

$$2 \sin mx \,.\, \cos nx = \sin (m+n)x + \sin (m-n)x,$$

and $$2 \cos mx \,.\, \cos nx = \cos (m+n)x + \cos (m-n)x,$$

and the results established in a manner similar to that employed for (3.5*a*). For (*d*),

$$\begin{aligned}\int_0^{2\pi} \sin kx \,.\, \cos kx\, dx &= \tfrac{1}{2}\int_0^{2\pi} \sin 2kx\, dx \\ &= \tfrac{1}{2}\left[\frac{-\cos 2kx}{2k}\right]_0^{2\pi}. \\ &= 0.\end{aligned}$$

For (*e*), $$\begin{aligned}\int_0^{2\pi} \sin kx\, dx &= \left[\frac{-\cos kx}{k}\right]_0^{2\pi}, \\ &= 0\end{aligned}$$

and similarly for (*f*). For (*h*), the result is used :

$$\cos 2kx = 2 \cos^2 kx - 1,$$

and the integration performed as for (*g*).

To derive the formulæ (3.3) for the coefficients in the series (3.2), first integrate both sides of this equation with respect to x over a cycle $(0 \leqslant x \leqslant 2\pi)$. Then as k is an integer, equations (*e*) and (*f*) of (3.5) give the result

$$\begin{aligned}\int_0^{2\pi} f(x)\, dx &= \int_0^{2\pi} \tfrac{1}{2}a_0\, dx \\ &= \pi a_0.\end{aligned}$$

Hence $$a_0 = \frac{1}{\pi}\int_0^{2\pi} f(x)\, dx. \quad . \quad . \quad . \quad \text{(i)}$$

Secondly, multiply (3.2) by cos kx and again integrate over a cycle. Then

$$\int_0^{2\pi} f(x)\cos kx\,dx = \int_0^{2\pi}\left\{\tfrac{1}{2}a_0 + \sum_{k=1}^{k=\infty}(a_k \cos kx + b_k \sin kx)\right\}\cos kx\,dx$$

$$= \pi a_k, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

by reason of equations (b), (c), (f) and (g) of (3.5). Similarly, it may be shown that

$$\int_0^{2\pi} f(x)\sin kx\,dx = \pi b_k. \quad . \quad . \quad . \quad \text{(iii)}$$

The equations (i)-(iii) are equivalent to (3.3), since putting $k = 0$ in the first equation of (3.3) gives the result (i). The whole process of determining the Fourier coefficients may be stated thus:

"The average value of the function over a cycle (i.e. from $x = 0$ to $x = 2\pi$) gives the constant term $\frac{1}{2}a_0$ (or A_0).

"The general cosine coefficient a_k is obtained by multiplying the function by the corresponding cosine function (cos kx) and taking twice the average of this product over a cycle.

"The general sine coefficient b_k is obtained by multiplying the function by the corresponding sine function (sin kx) and taking twice the average of this product over a cycle."

The procedure is now illustrated by an example.

Example.

The function represented by the line in Fig. 3a may be expressed in the form:

$$\left.\begin{aligned} y &= x/a \quad \text{for} \quad 0 \leqslant x < a \\ &= 0 \qquad\qquad a < x \leqslant 2\pi, \\ \text{and} \quad y &= f(x) = f(x + 2\pi) \\ &\text{for values of } x \text{ outside the interval } 0 - 2\pi. \end{aligned}\right\} \quad . \quad . \quad . \quad (3.7)$$

The coefficients are determined as under:

$$A_0 = \tfrac{1}{2}a_0 = \frac{1}{2\pi}\int_0^a \frac{x}{a}\,dx = \frac{1}{2\pi a}\left[\frac{x^2}{2}\right]_0^a = \frac{a}{4\pi}. \quad . \quad (3.8a)$$

The determination of the coefficients a_k and b_k depend upon the process of integration by parts. From the results of elementary differential calculus,

$$\frac{d}{dx}(uv) = u\frac{dv}{dx} + v\frac{du}{dx}.$$

Let $\qquad w = \dfrac{dv}{dx}$, so that $v = \int w\,dx$.

Then $$\frac{d}{dx}\left(u\int w\,dx\right) = uw + \frac{du}{dx}\int w\,dx,$$

and integration of this last equation with respect to x gives

$$u\int w\,dx = \int uw\,dx + \int\left(\frac{du}{dx}\int w\,dx\right)dx,$$

i.e. $$\int uw\,dx = u\int w\,dx - \int\left(\frac{du}{dx}\int w\,dx\right)dx.$$

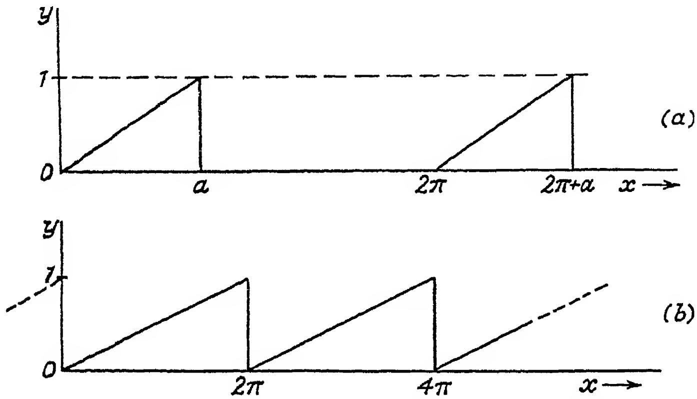


FIG. 3.—Periodic function expressed by (3.7), with coefficients as given by (3.8). At (b) the wave becomes the "saw-tooth wave" well-known to electronic engineers.

Applying this formula,

$$\int x\cos kx\,.\,dx = \frac{x}{k}\sin kx - \int\frac{\sin kx}{k}dx$$
$$= \frac{x}{k}\sin kx + \frac{\cos kx}{k^2},$$

and $$\int x\sin kx\,.\,dx = -\frac{x}{k}\cos kx - \int\frac{-\cos kx}{k}dx$$
$$= -\frac{x}{k}\cos kx + \frac{\sin kx}{k^2}.$$

Hence $$\pi a_k = \int_0^a \frac{x}{a}\cos kx\,dx$$
$$= \frac{1}{a}\left[\frac{x}{k}\sin kx + \frac{\cos kx}{k^2}\right]_0^a$$
$$= \frac{\sin ka}{k} - \frac{1 - \cos ka}{ak^2},$$

i.e. $$a_l = \frac{1}{k\pi}\left(\sin ka - \frac{1-\cos ka}{ak}\right) \quad . \quad (3.8b)$$

Similarly, it can be shown that

$$b_k = \frac{1}{k\pi}\left(-\cos ka + \frac{\sin ka}{ak}\right) \quad . \quad . \quad (3.8c)$$

For the special case where $a = 2\pi$ the waveform takes the form illustrated in Fig. 3b, and the coefficients have the values :

$$\left.\begin{aligned} A_0 &= \tfrac{1}{2}a_0 = \tfrac{1}{2} \\ a_k &= 0 \\ b_k &= -\frac{1}{k\pi} \end{aligned}\right\} \quad . \quad . \quad . \quad (3.9)$$

Figs. 4a, b, c show the synthetic waves produced by the addition of the first two, four, and six harmonics, respectively; i.e. the functions represented are :

Fig. 4a $y = \frac{1}{2} - \frac{1}{\pi}\left(\sin x + \frac{\sin 2x}{2}\right).$

Fig. 4b $y = \frac{1}{2} - \frac{1}{\pi}\left(\sin x + \frac{\sin 2x}{2} + \frac{\sin 3x}{3} + \frac{\sin 4x}{4}\right).$

Fig. 4c $y = \frac{1}{2} - \frac{1}{\pi}\left(\sin x + \frac{\sin 2x}{2} + \frac{\sin 3x}{3} + \frac{\sin 4x}{4} + \frac{\sin 5x}{5} + \frac{\sin 6x}{6}\right).$

The rapidity with which these curves approach the limiting curve (Fig. 3b) shows how accurately an arbitrary function can be represented by only a few harmonic terms. (See note in Section 7 on Gibbs' phenomenon.)

4. Half-range series (symmetry and skew-symmetry).

In Section 7 of Chapter II it was shown that a wave which is symmetrical about the values $x = 0$ and $x = \pi$ can be represented by a series containing cosine terms only, while a wave which is skew-symmetrical about these values can be represented by a series containing sine terms only; a constant term may be included in the symmetrical case, and the constant term involved in the case of disguised skew-symmetry may be removed, by a suitable change in the variable y, so as to leave a truly skew-symmetrical wave. Under conditions of symmetry or skew-symmetry the wave may

be represented by a so-called *half-range* Fourier series, such as will now be discussed.

Taking first the case of symmetry, let $F(x)$ be a periodic function with period 2π, symmetrical about the values $x = 0$ and $x = \pi$.

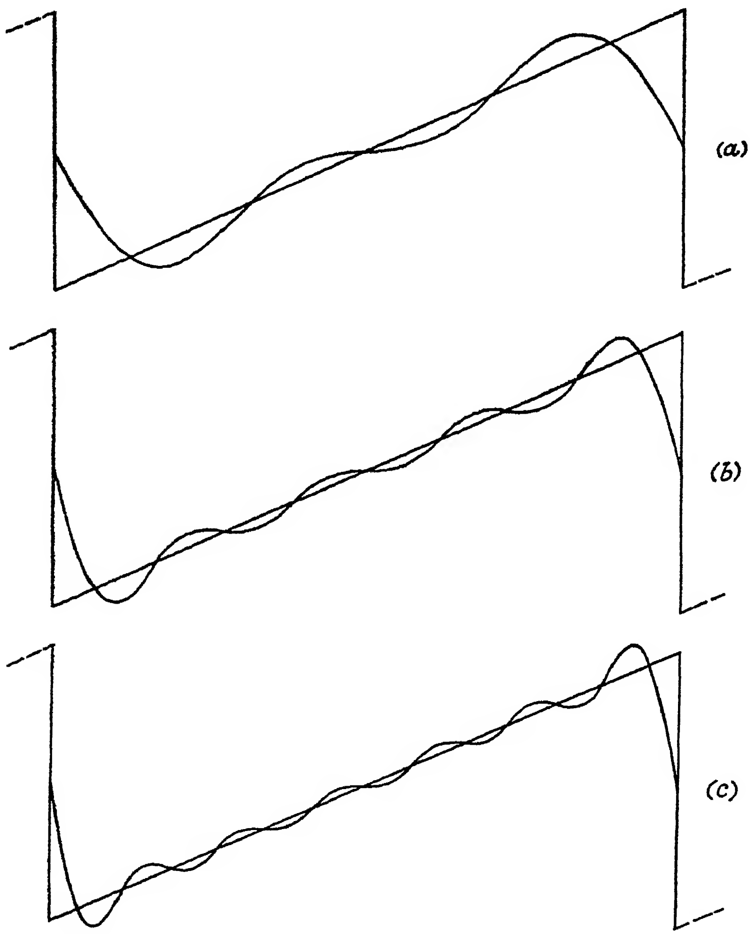


FIG. 4.—Sum of first two, four and six harmonics of the wave shown in Fig. 3*b*.

Then $$F(x) = F(2\pi - x),$$

or, if $$z = 2\pi - x, \quad F(x) = F(z).$$

Now, $$\int_0^{2\pi} F(x)\, dx = \int_0^{\pi} F(x)\, dx + \int_{\pi}^{2\pi} F(x)\, dx,$$

and, since $$dx = -dz,$$

$$\int_{\pi}^{2\pi} \mathrm{F}(x)\,dx = -\int_{\pi}^{0} \mathrm{F}(z)\,dz$$
$$= \int_{0}^{\pi} \mathrm{F}(z)\,dz = \int_{0}^{\pi} \mathrm{F}(x)\,dx,$$

Hence $$\int_{0}^{2\pi} \mathrm{F}(x)\,dx = 2\int_{0}^{\pi} \mathrm{F}(x)\,dx.$$

Now let $\mathrm{F}(x) = f(x)\cos kx$, where k is an integer. Then if $f(x)$ is a periodic function with a period 2π, which is symmetrical about the values $x = 0$ and $x = \pi$, $\cos kx$ fulfils the same conditions and so does $\mathrm{F}(x)$, so that

$$\int_{0}^{2\pi} f(x)\cos kx\,.\,dx = 2\int_{0}^{\pi} f(x)\cos kx\,.\,dx.$$

Hence
$$\left.\begin{aligned} \mathrm{A}_0 = \tfrac{1}{2}a_0 &= \frac{1}{\pi}\int_{0}^{\pi} f(x)\,dx, \\ a_k &= \frac{2}{\pi}\int_{0}^{\pi} f(x)\cos kx\,.\,dx. \end{aligned}\right\} \qquad (4.1)$$

Similarly, if $\mathrm{F}(x) = f(x)\sin kx$, where $f(x)$ is skew-symmetrical about the values $x = 0$ and $x = \pi$, then since $\sin kx$ fulfils these same conditions of skew-symmetry it follows that $\mathrm{F}(x)$ is symmetrical about these values; for

$$f(x) = -f(2\pi - x),$$
and
$$\sin kx = -\sin k(2\pi - x).$$

Hence
$$\mathrm{F}(x) = f(x)\sin kx = f(2\pi - x)\sin k(2\pi - x)$$
$$= \mathrm{F}(2\pi - x).$$

Thus in this case,

$$b_k = \frac{2}{\pi}\int_{0}^{\pi} f(x)\sin kx\,.\,dx. \qquad (4.2)$$

The series

$$y = \tfrac{1}{2}a_0 + \sum_{k=1}^{k=\infty} a_k \cos kx$$

representing a symmetrical wave-form, the coefficients being given by (4.1) and the series

$$y = \sum_{k=1}^{k=\infty} b_k \sin kx$$

representing a skew-symmetrical wave-form, the coefficients being given by (4.2), are termed *half-range series*, as the integration is

carried out over half a cycle only. Similarly, the series (3.2) representing a general waveform, the coefficients being given by (3.3), is termed a *full-range series*, as the integration is carried out over a full cycle.

TABLE I

Values of Fourier coefficients for square-peaked waveform (Fig. 5a)

See text for values of a greater than $\pi/2$ radians. Table gives values of general coefficient a_k for values of k from 0 to 12; constant term $A_0 = \frac{1}{2}a_0$.

a / k	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$
0	0·333	0·500	0·667	1
1	0·318	0·450	0·551	0·637
2	0·276	0·318	0·276	0
3	0·212	0·150	0	−0·212
4	0·138	0	−0·138	0
5	0·063	−0·090	−0·110	0·128
6	0	−0·106	0	0
7	−0·046	0·064	0·079	−0·091
8	−0·069	0	0·069	0
9	−0·071	0·050	0	0·071
10	−0·055	0·063	−0·055	0
11	0·029	0·041	0·050	−0·058
12	0	0	0	0

Examples.

By using the formulæ (4.1) and (4.2) whenever occasion permits the work involved in the evaluation of the integrals is considerably reduced when the function to be analysed is complicated. To illustrate the method some simple examples are first described.

(i) *Symmetry.*

Fig. 5a represents a "square-peaked" waveform, which is symmetrical about the points A and B. Let the axes of x and y be as shown in the diagram; then the function represented is

$$\left.\begin{aligned} y = 1 \quad &\text{for} \quad 0 \leqslant x < a, \\ &\text{and} \quad 2\pi - a < x \leqslant 2\pi, \\ \text{and} \qquad y = 0 \quad &\text{for} \quad a < x < 2\pi - a. \end{aligned}\right\} \quad . \quad . \quad (4.3)$$

From equations (4·1),

$$\tfrac{1}{2}a_0 = \frac{1}{\pi}\int_0^a dx = \frac{a}{\pi}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.4a)$$

$$a_k = \frac{2}{\pi}\int_0^a \cos kx\, dx = \frac{2}{k\pi}\Big[\sin kx\Big]_0^a$$

$$= \frac{2}{k\pi}\sin ka, \qquad . \quad . \quad . \quad . \quad . \quad . \quad (4.4b)$$

b_k is of course zero for all values of k, since the function is symmetrical and therefore contains no sine terms. The values of the coefficients (4.4) are clearly dependent upon the value of a; Table I gives the values of a_0 and the first twelve cosine coefficients for various typical values of a not greater than $\pi/2$. For values of

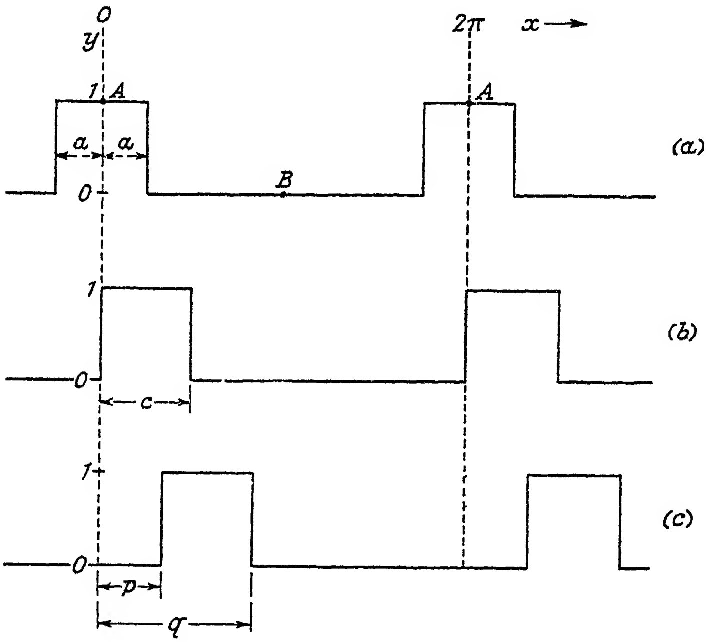


FIG. 5.—"Square-peaked" waveform.

a greater than $\pi/2$ radians, let $\pi - a = b$; then if the column headings in Table I are taken to represent b instead of a, the values of the *even* coefficients should have their signs changed, the *odd* harmonics remaining unaltered, while the constant term $A_0 = 1 - \frac{1}{2}a_0'$, where a_0' is obtained from the Table. Thus, for example, if $a = 2\pi/3$, $b = \pi/3$: $A_0 = \frac{1}{2}a_0 = 1 - 0{\cdot}333 = 0{\cdot}667$, $k_1 = 0{\cdot}551$, $k_2 = -0{\cdot}276$, etc.

The square-peaked waveform is of some importance in electrical theory, since the response of circuits is now being measured by reference to the transmission of such waveforms, furthermore,

these waveforms form the basis of the concept of *periodic existence functions* (see Section 6). For these reasons, they will now be further investigated.

The waveform illustrated in Fig. 5*b* is the same as that in Fig. 5*a* except that the function represented is now

$$\left.\begin{array}{lll} y = 1 & \text{for} & 0 < x < c \\ y = 0 & \text{for} & c < x < 2\pi \end{array}\right\} \qquad . \quad . \quad (4.5)$$

i.e., there has been a shift along the x-axis through a distance a, and $c = 2a$. Since any function when multiplied by this function will have zero value except over the range from 0 to c (and for corresponding parts of other cycles, since the function represented in Fig. 5 is periodic), and will have its own value over this range, the term "periodic existence function" is applied to the function (4.5). This concept is developed in Section 6, below. The Fourier coefficients of the series expansion of (4.5) may be obtained in two ways: by application of the basic formulæ (3.3), or by changing the variable x in the results (4.4).

From (3.3),

$$A_0 = \tfrac{1}{2}a_0 = \frac{1}{2\pi}\int_0^c dx = \frac{c}{2\pi}. \qquad . \quad . \quad . \quad . \quad (4.6a)$$

$$a_k = \frac{1}{\pi}\int_0^c \cos kx\, dx = \frac{1}{k\pi}\Big[\sin kx\Big]_0^c$$

$$= \frac{1}{k\pi}\sin kc. \qquad . \quad . \quad . \quad . \quad . \quad . \quad (4.6b)$$

$$b_k = \frac{1}{\pi}\int_0^c \sin kx\, dx = \frac{-1}{k\pi}\Big[\cos kx\Big]_0^c$$

$$= \frac{1}{k\pi}(1 - \cos kc). \qquad . \quad . \quad . \quad . \quad . \quad (4.6c)$$

Alternatively, if $x_1 = x - a$,

$$\cos kx = \cos(kx_1 - ka)$$

$$= \cos kx_1 \cos ka + \sin kx_1 \sin ka,$$

and the general term (the coefficient of which is given by (4.4*b*)) is

$$a_k \cos kx = \frac{2}{k\pi}\sin ka \,.\, \cos ka \,.\, \cos kx_1$$

$$+ \frac{2}{k\pi}\sin^2 ka \,.\, \sin kx_1$$

$$= \frac{1}{k\pi}\sin 2ka \,.\, \cos kx_1 + \frac{1}{k\pi}(1 - \cos 2ka)\sin kx_1.$$

Since $c = 2a$, this last equation gives the results (4.6*b*) and (4.6*c*). The constant term $A_0 = \frac{1}{2}a_0$ is of course unaltered by the shift along the x-axis.

In Fig. 5*c* the wave has been shifted again along the x-axis, so that the function represented is now

$$\left.\begin{array}{lll} y = 1 & \text{for} & p < x < q, \\ y = 0 & \text{for} & 0 \leqslant x < p, \\ & \text{and} & q < x \leqslant 2\pi \end{array}\right\} \quad . \quad . \quad (4.7)$$

Let the actual dimensions of the wave be the same as in Fig. 5*b*, so that $c = q - p$; then since $x = x_1 + p$, where x refers to Fig. 5*c* and x_1 to Fig. 5*b*,

$$\begin{aligned} \cos kx_1 &= \cos (kx - kp) \\ &= \cos kx \,.\, \cos kp + \sin kx \,.\, \sin kp, \end{aligned}$$

and

$$\begin{aligned} \sin kx_1 &= \sin (kx - kp) \\ &= \sin kx \,.\, \cos kp - \cos kx \,.\, \sin kp. \end{aligned}$$

The general term (i.e. representing the kth harmonic) from equations (4.6*b*) and (4.6*c*) is

$$\frac{1}{k\pi}[\sin kc \,.\, \cos kx_1 + (1 - \cos kc) \sin kx_1],$$

and omitting the multiplier $1/k\pi$ for convenience, this expression can be written as

$$\begin{aligned} &\sin k(q - p) \,.\, \cos kp \,.\, \cos kx_1 \\ &+ \sin k(q - p) \,.\, \sin kp \,.\, \sin kx_1 \\ &+ [1 - \cos k(q - p)] \cos kp \,.\, \sin kx_1 \\ &- [1 - \cos k(q - p)] \sin kp \,.\, \cos kx_1. \quad . \quad . \quad (4.8) \end{aligned}$$

By using the result (2.4) of Chapter I the following identities are easily derived:

$$\left.\begin{aligned} \sin (A + B) &= \sin A \,.\, \cos B + \cos A \,.\, \sin B \\ \sin (A - B) &= \sin A \,.\, \cos B - \cos A \,.\, \sin B \\ \cos (A + B) &= \cos A \,.\, \cos B - \sin A \,.\, \sin B \\ \cos (A - B) &= \cos A \,.\, \cos B + \sin A \,.\, \sin B \end{aligned}\right\} \quad . \quad (4.9)$$

Now (4.8) can be written in the form

$$\begin{aligned} &\cos kx_1 [\sin k(q - p) \,.\, \cos kp + \cos k(q - p) \,.\, \sin kp - \sin kp] \\ &\quad + \sin kx_1 [\sin k(q - p) \,.\, \sin kp - \cos k(q - p) \,.\, \cos kp + \cos kp], \end{aligned}$$

and this again, by using (4.9), can be written:

$$\cos kx_1 (\sin kq - \sin kp) - \sin kx_1 (\cos kp - \cos kq).$$

Hence the coefficients in the series representing (4.7) are

$$\left.\begin{aligned} A_0 &= \tfrac{1}{2}a_0 = \frac{q-p}{2\pi} \\ a_k &= \frac{1}{k\pi}(\sin kq - \sin kp) \\ b_k &= \frac{1}{k\pi}(\cos kp - \cos kq) \end{aligned}\right\} \quad . \quad . \quad . \quad (4.10)$$

On comparing (4.10) with (4.6) it can be seen that if the symbol E_p^q represents the function (4.7) then

$$E_p^q = E_0^q - E_0^p. \quad . \quad . \quad . \quad (4.11)$$

This equation states in effect that the range from p to q is the same as the difference between the ranges from 0 to q and from 0 to p, and may appear to be a truism; the formal statement (4.11), however, is useful in connection with the existence functions discussed in Section 6. The formulæ (4.10) could of course have been obtained directly from (3.3).

(ii) *Skew-symmetry.*

The function represented by Fig. 6*a* can be expressed:

$$\left.\begin{aligned} y &= 1 \quad \text{for} \quad 0 < x < a \\ y &= 0 \quad \text{for} \quad a < x < 2\pi - a \\ y &= -1 \quad \text{for} \quad 2\pi - a < x < 2\pi \end{aligned}\right\} \quad . \quad . \quad (4.12)$$

and is skew symmetrical about the points A and B. The formula (4.2) gives

$$\begin{aligned} b_k &= \frac{2}{\pi}\int_0^a \sin kx \,.\, dx = \frac{-2}{k\pi}\Big[\cos kx\Big]_0^a \\ &= \frac{2}{k\pi}(1 - \cos ka). \quad . \quad . \quad . \quad . \quad (4.13) \end{aligned}$$

There is no constant term and no cosine term in the series. If the function (4.12) is added to the function (4.3) then, denoting the sum by Y,

if $\quad 0 < x < a, \quad Y = 2,$
if $\quad a < x < 2\pi \quad Y = 0,$

and a comparison of equations (4.4) and (4.13) with (4.6) checks the evident result that (4.3) and (4.12) represent respectively twice

the symmetrical and skew-symmetrical parts of (4.5), with c replaced by a.

Table II gives the values of b_k ($k = 1, 2, \ldots 12$) for various values of a. When $a = \pi$, the waveform is as shown in Fig. 6b. This is the same form as is obtained if $a = \pi/2$ in (4.3), except that the basic variable has been increased by a quarter-period, and that the total amplitude of the function has been doubled. Comparison of the last columns of Tables I and II in the light of the results listed in Table II of Chapter II, relating to the quarter-period shift, confirms the agreement: the waveform is alternant, so that there are only

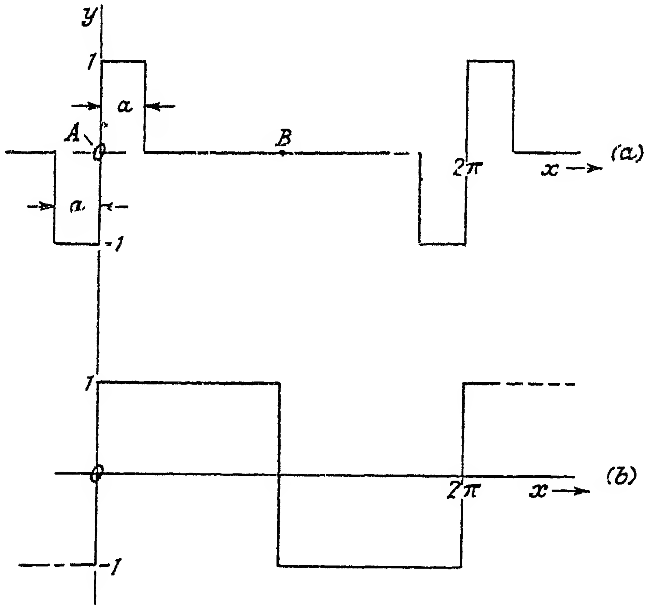


FIG. 6.—A skew-symmetrical function, which degenerates in a special case (b) to the same form as Fig. 5, with $c = \pi$.

odd harmonics present in the series, and the table in Chapter II shows that the first, fifth, etc., cosine terms become sine terms with the same sign, while the third, seventh, etc., cosine terms become sine terms with a change of sign.

With an alternant waveform (i.e. one for which the condition

$$f(x + \pi) = -f(x)$$

is true) which is both symmetrical and skew-symmetrical, it is frequently more convenient to make use of the property of skew-symmetry than that of symmetry in order to simplify the process of integration, as will now be shown by an example.

TABLE II

Values of Fourier coefficients for waveform illustrated in Fig. 6a

$a =$	$\pi/3$	$\pi/2$	$2\pi/3$	π
k				
1	0·318	0·637	0·955	1·274
2	0·477	0·637	0·477	0
3	0·424	0·212	0	0·424
4	0·239	0	0·239	0
5	0·064	0·127	0·191	0·255
6	0	0·212	0	0
7	0·045	0·091	0·136	0·182
8	0·119	0	0·119	0
9	0·141	0·071	0	0·141
10	0·096	0·127	0·096	0
11	0·029	0·058	0·087	0·116
12	0	0	0	0

The waveform illustrated in Fig. 7 consists of parabolic arcs; with the axes as shown, use is made of the property of skew-symmetry and the mathematical expression for the function formulated only for the range $0 \leqslant x \leqslant \pi$. Since $x = \pi/2$ gives the

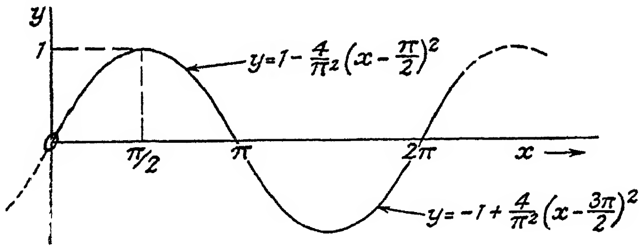


FIG. 7.—Approximation to sine wave by parabolic arcs. The coefficients of the harmonics are given in (4.17).

maximum value for y, and the curve is given as parabolic, over this range y is given by

$$y = 1 - a\left(x - \frac{\pi}{2}\right)^2. \quad . \quad . \quad . \quad . \quad (4.14)$$

where a has to be determined. Since $y = 0$ when $x = 0$, $a = 4/\pi^2$; checking for $x = \pi$, the corresponding value of y is given as 0, which is clearly correct. Substituting this value of a in (4.14),

$$y = 1 - \frac{4}{\pi^2}\left(x - \frac{\pi}{2}\right)^2$$

$$= -\frac{4x^2}{\pi^2} + \frac{4x}{\pi}. \quad . \quad . \quad . \quad . \quad . \quad (4.15)$$

In order to determine the Fourier coefficients by means of (4.2) it is first necessary to evaluate $\int x^2 \sin kx \,.\, dx$. This may be done in the same manner as that in which the results (3.8b) and (3.8c) were obtained:

$$\int x^2 \sin kx \,.\, dx = x^2\left(-\frac{\cos kx}{k}\right) - \int 2x\left(-\frac{\cos kx}{k}\right)dx$$

$$= -\frac{x^2}{k}\cos kx + \frac{2}{k}\int x \cos kx \,.\, dx$$

$$= -\frac{x^2}{k}\cos kx + \frac{2x}{k^2}\sin kx + \frac{2}{k^3}\cos kx.$$

Hence from (4.2) and (4.15), the coefficient of the general sine term in the series expansion of the function is given by

$$b_k = -\frac{8}{\pi^3}\int_0^\pi x^2 \sin kx \,.\, dx + \frac{8}{\pi^2}\int_0^\pi x \sin kx \,.\, dx$$

$$= -\frac{8}{\pi^3}\left[-\frac{x^2}{k}\cos kx + \frac{2x}{k^2}\sin kx + \frac{2}{k^3}\cos kx\right]_0^\pi$$

$$+ \frac{8}{\pi^2}\left[-\frac{x}{k}\cos kx + \frac{\sin kx}{k^2}\right]_0^\pi$$

$$= -\frac{8}{\pi^3}\left[\left(-\frac{\pi^2}{k} + \frac{2}{k^3}\right)\cos k\pi - \frac{2}{k^3}\right] + \frac{8}{\pi^2}\left[-\frac{\pi \cos k\pi}{k}\right]$$

$$= \frac{16}{k^3\pi^3}(1 - \cos k\pi). \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.16)$$

From (4.16),

$$\left.\begin{aligned} &\text{if } k \text{ is odd, } b_k = \frac{32}{k^3\pi^3} \\ &\text{if } k \text{ is even, } b_k = 0 \end{aligned}\right\} \quad . \quad . \quad . \quad (4.17)$$

The series representing the function (4.15) is therefore

$$y = \frac{32}{\pi^3}\left(\sin x + \frac{\sin 3x}{27} + \frac{\sin 5x}{125} + \text{etc.}\right)$$

and it is evident that a waveform composed of parabolic arcs in the manner of Fig. 7 is a very good approximation to a sine-wave, since the amplitude of the next term in the series is only one twenty-seventh of that of the fundamental, and all the other terms in the series are negligibly small.

If the origin of x is taken at a value giving the maximum value 1 to y (i.e. the wave is shifted along the x-axis through a distance representing $\pi/2$ radians in the negative direction), the function can be expressed in the form

$$\left.\begin{aligned} y &= 1 - ax^2 && \text{for } 0 \leqslant x \leqslant \pi/2, \\ y &= -1 + a(\pi - x)^2 && \text{for } \pi/2 \leqslant x \leqslant \pi, \end{aligned}\right\} \quad . \quad (4.18)$$

over the half range. If now the Fourier coefficients are evaluated from the formulæ (4.1) for a symmetrical waveform, the work involved in the integration is roughly double that involved in the determination of the Fourier coefficients from the formula (4.2) for a skew-symmetrical waveform. If the reader performs this evaluation for the coefficients of the symmetrical case, using the expressions (4.18), he will appreciate that much trouble can be saved by intelligent simplification of the problem in the manner indicated.

5. Further examples.

Two further general examples will be treated in this section.

(i) *Saw-tooth form.*

The waveform illustrated in Fig. 8*a* is the general saw-tooth form, taken as a skew-symmetrical function. The function can be expressed as

$$\left.\begin{aligned} y &= \frac{x}{a} && \text{for} \quad 0 < x < a \\ y &= 1 - \frac{x-a}{\pi-a} && \text{for} \quad a < x < \pi \\ &= \frac{\pi-x}{\pi-a} \end{aligned}\right\} \quad . \quad . \quad (5.1)$$

Since the waveform is skew-symmetrical about the values $x = 0$ and $x = \pi$, the series expansion contains sine terms only, and these can be evaluated by the formula (4.2). Thus :

$$\begin{aligned} \tfrac{1}{2}\pi b_k &= \frac{1}{a}\int_0^a x \sin kx \,.\, dx + \frac{1}{\pi-a}\int_a^\pi (\pi - x)\sin kx \,.\, dx \\ &= \frac{1}{a}\left[-\frac{x}{k}\cos kx + \frac{\sin kx}{k^2}\right]_0^a \\ &\qquad + \frac{1}{\pi-a}\left[-\frac{\pi}{k}\cos kx + \frac{x}{k}\cos kx - \frac{\sin kx}{k^2}\right]_a^\pi. \end{aligned}$$

This last equation reduces to

$$\tfrac{1}{2}\pi b_k = \frac{\pi}{k^2 a(\pi - a)} \sin ka,$$

and hence
$$b_k = \frac{2}{k^2 a(\pi - a)} \sin ka. \quad . \quad . \quad . \quad (5.2)$$

Three special cases arise. If $a = 0$, the waveform degenerates into that shown in Fig. 8*b*. The coefficients (5.2) must be evaluated

by L'Hospital's rule (3.6), since when $a = 0$, $\sin ka$ is also zero. Thus:

$$\underset{a \to 0}{\mathrm{Lt.}} \frac{\sin ka}{a} = \underset{a \to 0}{\mathrm{Lt.}} \frac{k \cos ka}{1}$$
$$= k,$$

and hence in this case

$$b_k = \frac{2}{k\pi}. \quad . \quad . \quad . \quad . \qquad (5.3)$$

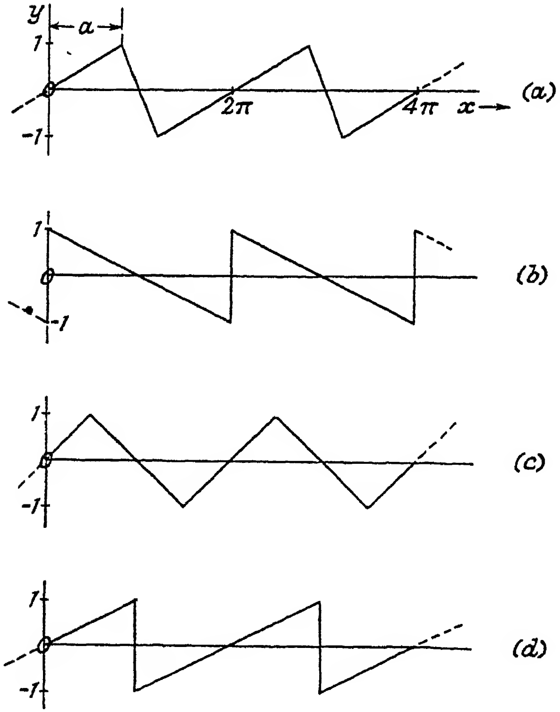


FIG. 8.—Saw-tooth waveforms. (b)-(d) are special cases of the general form (a).

If y represents the function (5.1), and y_1 represents the function (3.9) which is illustrated in Fig. 3b, it is evident that

$$y = 1 - 2y_1. \quad . \quad . \quad . \quad . \qquad (5.4)$$

Writing out the series in full,

$$y = \frac{2}{\pi}\left(\sin x + \frac{\sin 2x}{2} + \frac{\sin 3x}{3} + \text{etc.}\right),$$

and $$y_1 = \tfrac{1}{2} - \frac{1}{\pi}\left(\sin x + \frac{\sin 2x}{2} + \frac{\sin 3x}{3} + \text{etc.}\right),$$

and substitution of these expressions in (5.4) confirms the result. This procedure of comparing a special case of one waveform with that of another is very useful in checking the working.

If $a = \pi/2$ the resulting waveform is as shown in Fig. 8c. The coefficients are given by

$$b_k = \frac{2}{\frac{1}{4}k^2\pi^2}\sin k\pi/2$$
$$= \frac{8}{k^2\pi^2}\sin k\pi/2.$$

For even values of k, $\sin k\pi/2$ is zero; if $k = 1$, 5, 9, etc., $\sin k\pi/2 = 1$, while if $k = 3$, 7, 11, etc., $\sin k\pi/2 = -1$. Hence if

$$\left.\begin{array}{ll} k = 2n, & b_k = 0 \\ k = 4n + 1, & b_k = \dfrac{8}{k^2\pi^2} \\ k = 4n - 1, & b_k = \dfrac{-8}{k^2\pi^2} \end{array}\right\} \quad . \quad . \quad . \quad (5.5)$$

where n is an integer. That the even terms are absent from the series is evident from the fact that the wave is alternant (see Chapter II, Section 7).

If the origin of x is shifted to a value giving a maximum value to y, i.e. if the basic variable x is decreased by a quarter-period (or increased by three-quarters of a period), the resulting waveform is symmetrical about the values $x = 0$ and $x = \pi$. Utilising the results summarised in Table IV of Chapter II, the corresponding series is

$$y = \sum_{k=1}^{k=\infty} a_k \cos kx,$$

where if

$$\begin{array}{ll} k = 2n, & a_k = 0, \\ k = 4n + 1, & a_k = \dfrac{8}{k^2\pi^2}, \\ k = 4n - 1, & a_k = \dfrac{8}{k^2\pi^2}, \end{array}$$

i.e. if

$$\left.\begin{array}{ll} k = 2n, & a_k = 0 \\ k = 2n - 1, & a_k = \dfrac{8}{k^2\pi^2} \end{array}\right\} \quad . \quad . \quad (5.6)$$

where n is an integer.

When $a = \pi$, the waveform takes the form shown in Fig. 8*d*. Using L'Hospital's rule (3.6) to evaluate b_k from (5.2),

$$\underset{a \to \pi}{\text{Lt.}} \frac{\sin ka}{\pi - a} = \underset{a \to \pi}{\text{Lt.}} \frac{k \cos ka}{-1}$$

$$= -k \cos k\pi$$

Hence $$b_k = -\frac{2}{k\pi} \cos k\pi.$$

Now, for even values of k, $\cos k\pi = 1$, while for odd values of k, $\cos k\pi = -1$. Hence

$$\left.\begin{aligned} &\text{if } k \text{ is odd,} \quad b_k = \frac{2}{k\pi} \\ &\text{if } k \text{ is even,} \quad b_k = -\frac{2}{k\pi} \end{aligned}\right\} \qquad (5.7)$$

The graph (Fig. 8*d*) is the same as that in Fig. 8*b*, except that the wave has been shifted along the axis of x through a distance representing a half-period (i.e. the basic variable x has been increased by a half-period), and the sign of the function has been changed. Referring to Table III of Chapter II, it is seen that the effect of the half-period shift is to change the sign of the odd harmonic terms, while leaving the even harmonic terms unaltered. Comparison of the results (5.7) and (5.3) confirms this rule, it being remembered that the sign of the function has been changed in the transition from Fig. 8*b* to Fig. 8*d*.

A very close approximation to this type of waveform occurs in cathode-ray tube practice; a "saw-tooth" voltage is applied to one pair of deflector plates to give a relatively slow sweep in one direction, with a very rapid fly-back at the end of the stroke.

(ii) *Intermittent sine-wave.*

The waveform illustrated in Fig. 9*a* takes the form of half of a sine-wave cycle intermittently. Treating the wave as symmetrical, the function can be expressed as

$$\left.\begin{aligned} &y = \cos \frac{\pi x}{2a} \quad \text{for} \quad 0 \leqslant x \leqslant a \\ &y = 0 \qquad\quad\; \text{for} \quad a \leqslant x \leqslant \pi \end{aligned}\right\} \qquad (5.8)$$

Using the formulæ (4.1),

$$A_0 = \tfrac{1}{2}a_0 = \frac{1}{\pi}\int_0^a \cos Nx \,.\, dx,$$

where $\quad N = \dfrac{\pi}{2a}$. This integral gives

$$A_0 = \tfrac{1}{2}a_0 = \frac{1}{N\pi}\sin Na - \frac{1}{N\pi}\sin\frac{\pi}{2}$$

$$= \frac{1}{N\pi} - \frac{2a}{\pi^2}, \quad . \quad . \quad . \quad . \quad . \quad (5.9a)$$

and

$$a_k = \frac{2}{\pi}\int_0^a \cos Nx \,.\, \cos kx \,.\, dx$$

$$= \frac{1}{\pi}\int_0^a [\cos(N+k)x + \cos(N-k)x]dx$$

$$= \frac{1}{\pi}\left[\frac{\sin(N+k)x}{N+k} + \frac{\sin(N-k)x}{N-k}\right]_0^a,$$

i.e.

$$a_k = \frac{1}{\pi}\left(\frac{\sin(N+k)a}{N+k} + \frac{\sin(N-k)a}{N-k}\right). \quad . \quad (5.9b)$$

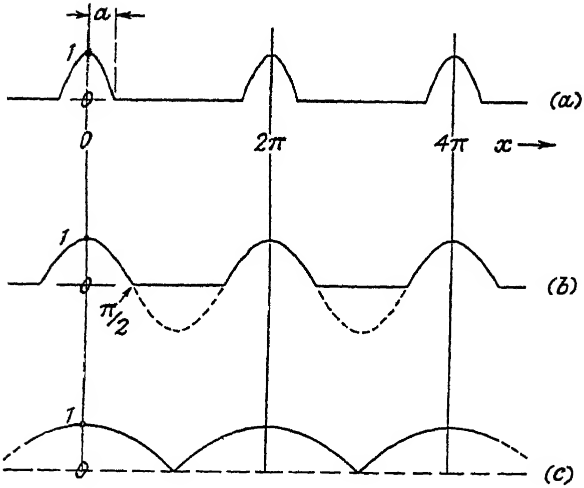


FIG. 9.—Intermittent sine wave and "rectified" waves. The half-rectified and full-rectified waves at (b) and (c) are special cases of the general form (a).

This expression can be put in an alternative form :

$$a_k = \frac{1}{\pi(N^2 - k^2)}[(N - k)\sin(N + k)a + (N + k)\sin(N - k)a]$$

$$= \frac{2}{\pi(N^2 - k^2)}(N\sin Na \,.\, \cos ka - k\cos Na \,.\, \sin ka).$$

But $\quad \mathrm{N}a = \frac{\pi}{2}$, hence $\sin \mathrm{N}a = 1$, $\cos \mathrm{N}a = 0$.

Therefore,
$$a_k = \frac{2\mathrm{N}}{\pi(\mathrm{N}^2 - k^2)} \cos ka. \quad \text{(5.9}c\text{)}$$

If N is integral, L'Hospital's rule (3.6) must be used to obtain the coefficient a_N. Since

$$\underset{\mathrm{N}\to k}{\mathrm{Lt.}} \frac{\sin(\mathrm{N}-k)a}{\mathrm{N}-k} = \underset{\mathrm{N}\to k}{\mathrm{Lt.}} \frac{a\cos(\mathrm{N}-k)a}{1}$$
$$= a,$$

and $\sin(\mathrm{N}+k)a = 0$ when $\mathrm{N} = k$, this coefficient is, from (5.9b),

$$a_\mathrm{N} = \frac{a}{\pi} = \frac{1}{2\mathrm{N}}. \quad \text{(5.9}d\text{)}$$

Two special cases of this form which are of particular importance are those in which $\mathrm{N} = 1$ and $\frac{1}{2}$. When $\mathrm{N} = 1$, $a = \pi/2$, and the wave has the form shown in Fig. 9b. This is the *half-rectified wave* familiar to electrical technicians. The coefficients are

$$\left.\begin{aligned} \mathrm{A}_0 = \tfrac{1}{2}a_0 &= \frac{1}{\pi} \\ a_1 &= \tfrac{1}{2} \\ \text{and, for} \quad k \neq 1,\ a_k &= 0 \qquad \text{if } k \text{ is odd} \\ &= \frac{-2}{\pi(k^2-1)} \qquad \text{if } k = 4n \\ &= \frac{2}{\pi(k^2-1)} \qquad \text{if } k = 4n+2 \end{aligned}\right\} \quad \text{(5.10)}$$

where n is an integer.

When $\mathrm{N} = \frac{1}{2}$, $a = \pi$, and the waveform has the form illustrated in Fig. 9c. This is the *full-rectified wave* well-known in electrical theory. The coefficients are:

$$\mathrm{A}_0 = \tfrac{1}{2}a_0 = \frac{2}{\pi}, \quad \text{(5.11}a\text{)}$$

$$a_k = \frac{-\cos k\pi}{\pi(k^2 - \frac{1}{4})}.$$

If k is odd, $\cos k = -1$, whereas if k is even, $\cos k = 1$. Hence

$$\left.\begin{aligned} a_k &= \frac{1}{\pi(k^2-\frac{1}{4})} \qquad \text{if } k \text{ is odd,} \\ &= \frac{-1}{\pi(k^2-\frac{1}{4})} \qquad \text{if } k \text{ is even} \end{aligned}\right\} \quad \text{(5.11}b\text{)}$$

The values of the constant term and the first twelve cosine coefficients for these two rectified waveforms are given in Table III. The actual values given are those of a_k; the constant term $A_0 = \frac{1}{2}a_0$, where a_0 is the value stated in the Table.

TABLE III

Values of Fourier coefficients a_k for the rectified waveforms (Figs. 9b and 9c). The constant term $A_0 = \frac{1}{2}a_0$.

k	Half-rectified (Fig 9b)	Full-rectified (Fig 9c)
0	0·637	1·274
1	0·500	0·425
2	0·212	−0·085
3	0	0·036
4	−0·042	−0·020
5	0	0·013
6	0·018	−0·009
7	0	0·007
8	−0·010	−0·005
9	0	0·004
10	0·006	−0·003
11	0	0·003
12	−0·004	−0·002

It is to be noted that the amplitudes of the higher even harmonics of the full-rectified wave are half those of the half-rectified wave.

The general form (5.8) is of importance in many practical investigations, for example, in the determination of the periodic forces imposed on a rotating aero-propeller by disturbances in the airflow, which may be caused by such obstructions as radiators, etc.

6. Periodic existence functions, and derived forms.

The function illustrated in Fig. 10*a*, which may be expressed as:

$$\left.\begin{aligned} y &= 1 \quad \text{for} \quad p < x < q \\ y &= 0 \quad \text{for} \quad 0 < x < p \\ &\qquad \text{and} \quad q < x < 2\pi \\ \text{and} \qquad y &= f(x) = f(2\pi + x) \end{aligned}\right\} \quad . \quad . \quad (6.1)$$

is termed a *periodic existence function with the interval* (p, q). The function may be said to exist, i.e. to have a non-zero value, over this interval; hence the name.

The notation E_p^q will be used to denote the function (6.1). It

is to be noted that the function is periodic, the period being 2π radians, so that by definition

$$E_p^q = E_{2\pi+p}^{2\pi+q}.$$

From a consideration of the diagrams in Figs. 10*a*, *b*, *c*, which represent respectively the functions

$$E_p^q,\ E_0^q,\ \text{and}\ E_0^p,$$

it is clear that the relation holds:

$$E_p^q = E_0^q - E_0^p. \quad . \quad . \quad . \quad . \quad (6.2)$$

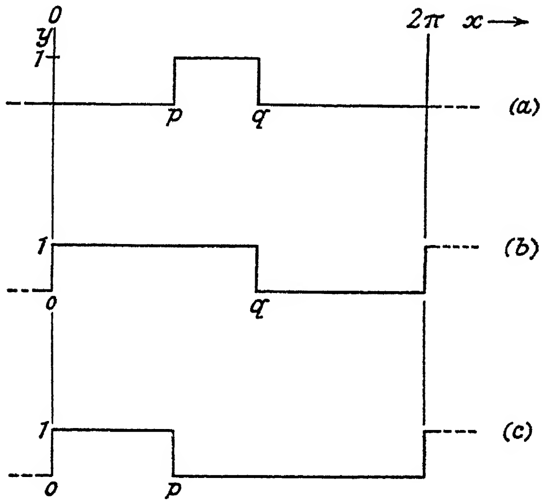


FIG. 10.—Periodic existence functions. The function (*a*), operative over the interval *pq*, is the difference between the functions (*b*) and (*c*).

The function E_p^q can be expressed in the form of a Fourier series; in fact, this expansion has already been given in Section 4 of this chapter. The formulæ (4.10) for the coefficients of this series are here repeated for ease of reference:

$$\left.\begin{aligned} A_0 &= \tfrac{1}{2}a_0 = \frac{q-p}{2\pi} \\ a_k &= \frac{1}{k\pi}(\sin kq - \sin kp) \\ b_k &= \frac{1}{k\pi}(\cos kp - \cos kq) \end{aligned}\right\} \quad . \quad . \quad . \quad (6.3)$$

The symbols $(A_0)_p^q$, $(a_k)_p^q$, and $(b_k)_p^q$ will be used here to denote the coefficients (6.3) of the function E_p^q.

The periodic existence functions are useful when it is required to express a function $F(x)$, which is defined by a number of equations of the form

$$\left.\begin{aligned} F(x) &= f_0^a(x) \quad \text{for} \quad 0 < x < a \\ F(x) &= f_a^b(x) \quad \text{for} \quad a < x < b \\ &\text{etc., etc., } \ldots \\ F(x) &= f_n^{2\pi}(x) \quad \text{for} \quad n < x < 2\pi \end{aligned}\right\} \quad . \quad . \quad (6.4)$$

in a single formula. For since multiplication of a function by an existence function is equivalent to limiting its range to the interval of the existence function, i.e.

$$\begin{aligned} E_p^q \cdot f(x) &= f(x) \quad \text{for} \quad p < x < q \\ &= 0 \qquad \text{for} \quad 0 < x < p, \\ &\qquad\quad \text{and} \quad q < x < 2\pi, \end{aligned}$$

the function $F(x)$ can be represented over the whole range $(0, 2\pi)$ by the sum

$$F(x) = E_0^a \cdot f_0^a(x) + E_a^b \cdot f_a^b(x) + \ldots + E_n^{2\pi} \cdot f_n^{2\pi}(x), \quad . \quad (6.5)$$

and if the coefficients of $F(x)$, analogous to those of a Fourier series, are $\bar{A}_0$, $\bar{a}_k$, $\bar{b}_k$, then

$$\left.\begin{aligned} \bar{A}_0 &= (A_0)_0^a \cdot f_0^a(x) + \ldots + (A_0)_n^{2\pi} \cdot f_n^{2\pi}(x) \\ \bar{a}_k &= (a_k)_0^a \cdot f_0^a(x) + \ldots + (a_k)_n^{2\pi} \cdot f_n^{2\pi}(x) \\ \bar{b}_k &= (b_k)_0^a \cdot f_0^a(x) + \ldots + (b_k)_n^{2\pi} \cdot f_n^{2\pi}(x) \end{aligned}\right\} \quad . \quad (6.6)$$

It must be noted, however, that unless all the functions $f_0^a(x)$, etc., are constants, the series

$$F(x) = \bar{A}_0 + \sum_{k=1}^{k=\infty} \bar{a}_k \cos kx + \sum_{k=1}^{k=\infty} \bar{b}_k \sin kx$$

is not a Fourier series, since the coefficients are themselves functions of x. If the functions are all constants, however, the series is a Fourier series.

Two special forms of existence function apply to the cases of symmetrical and skew-symmetrical waveforms.

(i) *Symmetry*.

Let $\qquad (E')_p^q = E_p^q + E_{2\pi-q}^{2\pi-p} \; (0 \leqslant p < q \leqslant \pi). \quad . \quad . \quad (6.7)$

This function (6.7) is then symmetrical about the values $x = 0$

and $x = \pi$. If A_0, a_l, and b_k are the Fourier coefficients of the function $E_{2\pi-q}^{2\pi-p}$, then

$$A_0 = \frac{1}{2\pi}(2\pi - p - 2\pi + q) = \frac{1}{2\pi}(q - p) = (A_0)_p^q,$$

$$a_l = \frac{1}{k\pi}[\sin k(2\pi - p) - \sin k(2\pi - q)]$$

$$= \frac{-1}{k\pi}(\sin kp - \sin kq) = (a_k)_p^q,$$

and $$b_l = \frac{1}{k\pi}[\cos k(2\pi - q) - \cos k(2\pi - p)]$$

$$= \frac{1}{k\pi}(\cos kq - \cos kp) = -(b_k)_p^q,$$

and hence if $(A_0')_p^q$, $(a_k')_p^q$, $(b_k')_p^q$ are the Fourier coefficients of the function (6.7), then

$$\left.\begin{aligned} (A_0')_p^q &= 2(A_0)_p^q \\ (a_k')_p^q &= 2(a_k)_p^q \\ (b_k')_p^q &= 0 \end{aligned}\right\} \quad . \quad . \quad . \quad (6.8)$$

(ii) *Skew-symmetry.*

Let $$(E'')_p^q = E_p^q - E_{2\pi-q}^{2\pi-p} \quad . \quad . \quad . \quad (6.9)$$

which is skew-symmetrical about the values $x = 0$ and $x = \pi$; if

$$(A_0'')_p^q, \ (a_k'')_p^q, \ (b_k'')_p^q$$

are the Fourier coefficients of (6.9), then

$$\left.\begin{aligned} (A_0'')_p^q &= 0 \\ (a_k'')_p^q &= 0 \\ (b_k'')_p^q &= 2(b_k)_p^q \end{aligned}\right\} \quad . \quad . \quad . \quad (6.10)$$

Linear function. A similar type of procedure is useful in the analysis of waveforms made up of linear functions. Suppose that a function $y = f(x)$ is defined as follows:

$$\left.\begin{aligned} y &= c + mx \quad \text{for} \quad p < x < q \\ &= 0 \qquad\quad\ \ \text{for} \quad 0 < x < p \\ &\qquad\qquad\quad \text{and} \quad q < x < 2\pi \\ \text{and} \quad f(x) &= f(x + 2\pi). \end{aligned}\right\} \quad . \quad . \quad (6.11)$$

The Fourier coefficients are given by the expressions

$$A_0 = \tfrac{1}{2}a_0 = \frac{1}{2\pi}\int_p^q (c + mx)dx = \frac{1}{2\pi}\Big[cx + \tfrac{1}{2}mx^2\Big]_p^q$$

$$= \frac{1}{2\pi}[c(q - p) + \tfrac{1}{2}m(q^2 - p^2)]$$

$$= \frac{q - p}{2\pi}[c + \tfrac{1}{2}m(q + p)], \quad . \quad . \quad . \quad . \quad (6.12a)$$

$$a_k = \frac{1}{\pi}\int_p^q (c + mx) \cos kx \,.\, dx$$

$$= \frac{1}{\pi}\Big[\frac{c}{k} \sin kx + \frac{mx}{k} \sin kx + \frac{m}{k^2} \cos kx\Big]_p^q$$

$$= \frac{1}{k\pi}\Big[(c + mq) \sin kq - (c + mp) \sin kp$$

$$+ \frac{m}{k}(\cos kq - \cos kp)\Big], \quad (6.12b)$$

and $$b_k = \frac{1}{\pi}\int_p^q (c + mx) \sin kx \,.\, dx$$

$$= \frac{1}{\pi}\Big[- \frac{c}{k} \cos kx - \frac{mx}{k} \cos kx + \frac{m}{k^2} \sin kx\Big]_p^q$$

$$= \frac{1}{k\pi}\Big[(c + mp) \cos kp - (c + mq) \cos kq$$

$$+ \frac{m}{k}(\sin kq - \sin kp)\Big]. \quad (6.12c)$$

These formulæ are included in this work so that the analysis of certain types of function may be facilitated. In order to eliminate, for the research worker, much of the "donkey work" involved in the evaluation of the Fourier coefficients of arbitrarily defined functions, some indefinite integrals are listed below.

Indefinite integrals.

$$\int x \sin kx \,.\, dx = x\left(\frac{- \cos kx}{k}\right) - \int \frac{- \cos kx}{k} dx$$

$$= - \frac{x}{k} \cos kx + \frac{1}{k^2} \sin kx. \quad . \quad . \quad (6.13a)$$

$$\int x \cos kx \,.\, dx = x\left(\frac{\sin kx}{k}\right) - \int \frac{\sin kx}{k} dx$$

$$= \frac{x}{k} \sin kx + \frac{1}{k^2} \cos kx. \quad . \quad . \quad . \quad (6.13b)$$

$$\int x^2 \sin kx \,.\, dx = x^2\left(\frac{-\cos kx}{k}\right) - \int 2x\left(\frac{-\cos kx}{k}\right) dx$$

$$= -\frac{x^2}{k}\cos kx + \frac{2x}{k^2}\sin kx + \frac{2}{k^3}\cos kx. \quad (6.13c)$$

$$\int x^2 \cos kx \,.\, dx = x^2\left(\frac{\sin kx}{k}\right) - \int 2x\left(\frac{\sin kx}{k}\right) dx$$

$$= \frac{x^2}{k}\sin kx + \frac{2x}{k^2}\cos kx - \frac{2}{k^3}\sin kx. \quad (6.13d)$$

$$\int x^3 \sin kx \,.\, dx = x^3\left(\frac{-\cos kx}{k}\right) - \int 3x^2\left(\frac{-\cos kx}{k}\right) dx$$

$$= -\frac{x^3}{k}\cos kx + \frac{3x^2}{k^2}\sin kx + \frac{6x}{k^3}\cos kx$$

$$- \frac{6}{k^4}\sin kx. \quad (6.13e)$$

$$\int x^3 \cos kx \,.\, dx = x^3\left(\frac{\sin kx}{k}\right) - \int 3x^2\left(\frac{\sin kx}{k}\right) dx$$

$$= \frac{x^3}{k}\sin kx + \frac{3x^2}{k^2}\cos kx - \frac{6x}{k^3}\sin kx$$

$$- \frac{6}{k^4}\cos kx. \quad (6.13f)$$

$$\int \sin \mathrm{N}x \,.\, \sin kx \,.\, dx = \tfrac{1}{2}\int \cos (\mathrm{N}-k)x \,.\, dx - \tfrac{1}{2}\int \cos (\mathrm{N}+k)x \,.\, dx$$

$$= \tfrac{1}{2}\left[\frac{\sin (\mathrm{N}-k)x}{\mathrm{N}-k} - \frac{\sin (\mathrm{N}+k)x}{\mathrm{N}+k}\right]. \quad (6.13g)$$

$$\int \sin \mathrm{N}x \,.\, \cos kx \,.\, dx = \tfrac{1}{2}\int \sin (\mathrm{N}+k)x \,.\, dx + \tfrac{1}{2}\int \sin (\mathrm{N}-k)x \,.\, dx$$

$$= -\tfrac{1}{2}\left[\frac{\cos (\mathrm{N}+k)x}{\mathrm{N}+k} + \frac{\cos (\mathrm{N}-k)x}{\mathrm{N}-k}\right]. \quad (6.13h)$$

$$\int \cos \mathrm{N}x \,.\, \sin kx \,.\, dx = \tfrac{1}{2}\int \sin (\mathrm{N}+k)x \,.\, dx - \tfrac{1}{2}\int \sin (\mathrm{N}-k)x \,.\, dx$$

$$= \tfrac{1}{2}\left[\frac{\cos (\mathrm{N}-k)x}{\mathrm{N}-k} - \frac{\cos (\mathrm{N}+k)x}{\mathrm{N}+k}\right]. \quad (6.13i)$$

$$\int \cos \mathrm{N}x \,.\, \cos kx \,.\, dx = \tfrac{1}{2}\int \cos (\mathrm{N}+k)x \,.\, dx + \tfrac{1}{2}\int \cos (\mathrm{N}-k)x \,.\, dx$$

$$= \tfrac{1}{2}\left[\frac{\sin (\mathrm{N}+k)x}{\mathrm{N}+k} + \frac{\sin (\mathrm{N}-k)x}{\mathrm{N}-k}\right]. \quad (6.13j)$$

These expressions, in conjunction with the existence functions, will enable the Fourier coefficients of an arbitrarily defined periodic function to be evaluated speedily if the separate functions of which it is composed are constants, linear, quadratic or cubic functions of x, or any of these plus a trigonometric function of the type $\sin Nx$ or $\cos Nx$. All that needs to be done is the insertion of limits in the (at present) indefinite integrals (6.13) for each part of the arbitrary function, and the Fourier series so determined from the formulæ (3.3) can then be added directly to obtain the series expansion of the function.

7. Convergence of Fourier series ; Gibbs' phenomenon.

It is outside the scope of this book to consider the convergence of Fourier series in a rigorous mathematical manner ; for such discussions reference should be made to any textbook on Fourier series or mathematical analysis (see, for example, references 5, 6 and 7 in the Bibliography at the back of the book). In this section the subject will be treated as fully as is necessary for practical purposes. So long as the main principles involved are borne in mind it is unnecessary for the technician to be fully conversant with the mathematical niceties of the subject.

Determination of coefficients by consideration of least squared error. Let $f(x)$ be the function, defined over the period from $x = 0$ to $x = 2\pi$ radians, to which it is required to fit a Fourier series. Let S_n represent the sum of this series as far as the nth harmonics,

i.e.
$$S_n = A_0 + \sum_{k=1}^{k=n} (a_k \cos kx + b_k \sin kx). \quad . \quad . \quad (7.1)$$

It is required to find the values of the coefficients which will give the nearest approximation S_n to $f(x)$. First it is necessary to prescribe some means of assessing the nearness of an approximation. For this purpose the theory of the *mean squared error* is applied. Denoting the mean squared error by E, then

$$E = \frac{1}{2\pi}\int_0^{2\pi} [f(x) - S_n]^2 dx, \quad . \quad . \quad . \quad (7.2)$$

and it is seen that E is in fact a measure of the total deviation of the series S_n from the function $f(x)$; the deviation is squared so as to avoid the possibility of two deviations of opposite sign "cancelling out." The most accurate approximation will be obtained when E is a minimum, i.e.

$$\frac{\partial E}{\partial a_k} = \frac{\partial E}{\partial b_k} = \frac{\partial E}{\partial A_0} = 0. \quad . \quad . \quad . \quad (7.3)$$

The meaning of this equation (7.3) is as follows. It is assumed that if all the coefficients A_0, a_k, b_k are kept constant except one, say a_j, then as a_j is given different values the error E will vary; and it may be possible to determine a value of a_j which gives a minimum value to E. a_j is now fixed at such an optimum value, and some other coefficient, say a_t, varied. The procedure is repeated for all the coefficients, and is summarised in the equations (7.3). It is assumed that when all the coefficients have been found in this manner, the corresponding series S_n is the best approximation, as far as the nth harmonic, that can be found.

Substituting for E from (7.2) in (7.3), and solving first for a_k,

$$\frac{\partial E}{\partial a_k} = -\frac{1}{\pi}\int_0^{2\pi}[f(x) - S_n]\frac{\partial S_n}{\partial a_k}\,dx = 0.$$

Hence, $$\int_0^{2\pi}[f(x) - S_n]\cos kx\,.\,dx = 0,$$

and $$\int_0^{2\pi} f(x)\cos kx\,.\,dx = \int_0^{2\pi} a_k \cos^2 kx\,.\,dx,$$

by reason of equations (3.5), which give

$$\int_0^{2\pi} S_n \cos kx\,.\,dx = \int_0^{2\pi} a_k \cos^2 kx\,.\,dx$$
$$= \pi a_k.$$

Hence, $$a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\,.\,dx. \qquad (7.4a)$$

In a precisely similar manner the values of b_k and A_0 are found to be

$$b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin kx\,.\,dx, \qquad (7.4b)$$

$$A_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,.\,dx, \qquad (7.4c)$$

and these equations (7.4) are exactly the same as (3.3). All that has been proved so far is that the values (7.4) for the coefficients give a series S_n which is either the most accurate approximation to that number of terms (i.e. as far as the nth harmonic) or the least accurate approximation. It has not yet been determined whether the approximation is best or worst, or whether there is any advantage in taking more terms in the series.

On consideration it is evident that the approximation is best rather than worst, i.e. that the value of E corresponding to the determined series S_n is a minimum and not a maximum; for if $f(x)$ is finite throughout the cycle, as it will be in all practical cases,

the coefficients determined by (7.4) are all finite, and clearly a worse approximation can be found by making all the coefficients very large. Hence if the number n of harmonics included in the series S_n is fixed, the coefficients determined by (7.4) give the best approximation to the function $f(x)$. It remains to determine whether there is any advantage in taking a larger number of terms in the series, which is the informal method of determining whether the series (3.2) is convergent.

Let the mean squared error for the sum S_n be E_n, and that for S_{n+1}, E_{n+1}. Now since

$$E_n = \frac{1}{2\pi}\int_0^{2\pi} [f(x) - S_n]^2 dx,$$

and $$S_{n+1} = S_n + a_{n+1} \cos (n+1)x + b_{n+1} \sin (n+1)x,$$

$$\begin{aligned} E_{n+1} &= \frac{1}{2\pi}\int_0^{2\pi} [f(x) - S_n - a_{n+1} \cos (n+1)x - b_{n+1} \sin (n+1)x]^2 dx \\ &= E_n - \frac{1}{\pi}\int_0^{2\pi} [f(x) - S_n][a_{n+1} \cos (n+1)x + b_{n+1} \sin (n+1)x] dx \\ &\quad + \frac{1}{2\pi}\int_0^{2\pi} [a_{n+1} \cos (n+1)x + b_{n+1} \sin (n+1)x]^2 dx. \end{aligned}$$

Now the first integral term in this last equation is equal to

$$-\frac{1}{\pi}\int_0^{2\pi} f(x)[a_{n+1} \cos (n+1)x + b_{n+1} \sin (n+1)x] dx$$
$$= -(a_{n+1}^2 + b_{n+1}^2),$$

since there are no terms involving sine or cosine functions of $(n+1)x$ in S_n, and the second integral term is equal to

$$\tfrac{1}{2}(a_{n+1}^2 + b_{n+1}^2),$$

by reason of equations (3.5d, g, h). Hence the result is obtained:

$$E_{n+1} = E_n - \tfrac{1}{2}(a_{n+1}^2 + b_{n+1}^2). \qquad . \qquad . \qquad (7.5)$$

Equation (7·5) shows that the mean squared error when the first $(n+1)$ harmonics are included cannot be greater than the error when the first n harmonics are included, and if the $(n+1)$th harmonic is not zero the error is decreased by its inclusion. Hence the more terms are taken in the Fourier series, the more accurately will the function $f(x)$ be represented.

It still remains to demonstrate that as more and more terms are taken the error converges to zero, i.e. that

$$\underset{n\to\infty}{\text{Lt.}}\, E_n = 0.$$

Writing E_0 for the error when the constant term A_0 alone is taken in the series,

$$E_0 = \frac{1}{2\pi}\int_0^{2\pi}[f(x) - A_0]^2 dx$$
$$= \frac{1}{2\pi}\int_0^{2\pi} f(x)^2 dx - A_0^2,$$

and hence

$$E_1 = \frac{1}{2\pi}\int_0^{2\pi} f(x)^2 dx - A_0^2 - \tfrac{1}{2}(a_1^2 + b_1^2).$$

Proceeding in this manner, it is found that

$$E_n = \frac{1}{2\pi}\int_0^{2\pi} f(x)^2 dx - A_0^2 - \tfrac{1}{2}\sum_{k=1}^{k=n}(a_k^2 + b_k^2). \quad . \quad (7.6)$$

It is a well-known theorem in the analysis of infinite series that, if $f(0) = f(2\pi)$, and $f(x)$ is continuous in the interval from 0 to 2π,

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)^2 dx = A_0^2 + \tfrac{1}{2}\sum_{k=1}^{k=\infty}(a_k^2 + b_k^2)$$

(for a proof of this result, known as the Hurwitz-Liapounoff theorem, see reference 8 in the Bibliography). Hence as n is made greater and greater, the error E_n converges to zero.

The practical result of this analysis is that if the arbitrary function $f(x)$ is continuous within the interval of the cycle, the Fourier series can be made to approximate the function as closely as is desired by making n large enough, i.e. by taking a sufficient number of terms.

Gibbs' phenomenon.

When the arbitrary function has finite discontinuities, however, as in the case of those functions illustrated in Figs. 3, 5, 6, 8*b*, 8*d*, and 10, this result is not true. The function

$$\left.\begin{aligned} f(x) &= -\frac{\pi}{2} \quad \text{for} \quad \pi < x < 2\pi \\ &= \frac{\pi}{2} \quad \text{for} \quad 0 < x < \pi \\ \text{and} \quad f(0) &= f(\pi) = f(2\pi) = 0 \end{aligned}\right\} \quad . \quad . \quad (7.7)$$

is illustrated in Fig. 11*a*, and can be represented by the Fourier series

$$f(x) = 2\left(\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \ldots\right). \quad . \quad (7.8)$$

Let $$S_n(x) = 2\sum_{k=1}^{k=n}\frac{\sin(2k-1)x}{2k-1},$$

which is the series (7.8) taken as far as the $(2n - 1)$th harmonic. Then

$$\frac{d}{dx}S_n(x) = 2\sum_{k=1}^{k=n} \cos\,(2k-1)x = \frac{\sin 2nx}{\sin x},$$

and hence $$S_n(x) = \int_0^x \frac{\sin 2nt}{\sin t}\,dt.$$

Thus $$S_n(x) - \int_0^x \frac{\sin 2nt}{t}\,dt = \int_0^x (\sin 2nt)\left(\frac{1}{\sin t} - \frac{1}{t}\right)dt.$$

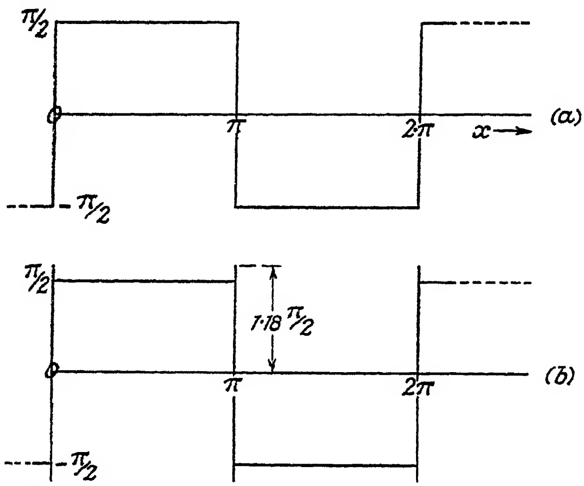


FIG. 11. Gibbs' phenomenon. The Fourier series for the function illustrated at (a) converges to the form (b), where the finite "jumps" are exceeded by about 9 per cent.

Now $$\int_0^x \frac{\sin 2nt}{t}\,dt = \int_0^{2nx} \frac{\sin t}{t}\,dt,\text{*}$$

and hence

$$S_n(x) - \int_0^{2nx} \frac{\sin t}{t}\,dt = \int_0^x (\sin 2nt)\,\frac{t}{\sin t}\left(\frac{t}{3\,!} - \frac{t^3}{5\,!} + \text{etc.}\right)dt.$$

Now as x continually increases from 0 to $\pi/2$, x . cosec x continually increases from 1 to $\pi/2$, and

$$0 < \frac{x}{3\,!} - \frac{x^3}{5\,!} + \text{etc.} < \frac{x}{3\,!}.$$

Thus $$\Delta = \left|S_n(x) - \int_0^{2nx} \frac{\sin t}{t}\,dt\right| < \frac{\pi}{12}\int_0^x x\,dx$$

$$< \frac{\pi}{24}x^2 \quad \text{when} \quad 0 < x \leqslant \pi/2.$$

* See section 9, page 181, for a proof of this result.

No matter how small the arbitrary positive quantity ϵ is chosen, a positive quantity m can be found such that

$$\Delta < \epsilon \quad \text{when} \quad 0 < x < m$$

for all values of n. If n be chosen so large that

$$\frac{\pi}{2n} < m,$$

then
$$\left| S_n\left(\frac{\pi}{2n}\right) - \int_0^{\pi} \frac{\sin t}{t}\, dt \right| < \epsilon. \qquad (7.9)$$

Now
$$\int_0^{\infty} \frac{\sin t}{t}\, dt = \frac{\pi}{2},$$

as is shown in section 9, page 182, and

$$\int_0^{\pi} \phi(t)\, dt = \int_0^{\infty} \phi(t)\, dt - \int_{\pi}^{\infty} \phi(t)\, dt,$$

hence from (7.9) the approximation curve $y = S_n(x)$ exceeds the value $\pi/2$ in the immediate positive neighbourhood of the origin by the amount

$$-\int_{\pi}^{\infty} \frac{\sin t}{t}\, dt = 0{\cdot}281.$$

From considerations of symmetry it is evident that at the finite discontinuities of the function the series overshoots the jump by about 9 per cent. of the jump, as illustrated in Fig. 11*b*.

This phenomenon is known as "Gibbs' phenomenon," for a full description of which see reference 9 in the Bibliography; the discussion given above is taken substantially from this work. The practical significance of the phenomenon is that the Fourier series approximation to a discontinuous function is not very reliable in the neighbourhood of the discontinuities unless a very large number of terms is taken; and even when many terms are included the series does not actually approximate to the function in these regions, but to a composite function made up of the original function plus the "overshoot" at either end of the line of discontinuity.

Rapidity of convergence of series. At parts of the wave where there is no discontinuity the Fourier series converges to the correct values; but if there are discontinuities anywhere in the cycle it is inherent in the nature of the series that the convergence in the continuous portions of the wave will be slow—i.e. that many terms will have to be included before a reasonable approximation is obtained. Fig. 12 shows, on a greatly magnified scale, part of the curves given by taking the first 5, 7, 9, 11, and 13 terms in the

series (7.8). Even when 13 terms are taken there is considerable excursion from the mean position indicated by the broken line.

An interesting general property concerning the rapidity of convergence of the Fourier series representing a periodic function has been mentioned by Hussman (reference 10 in the Bibliography at the end of the book). This may be stated as follows: the value of the kth Fourier coefficient is in general expressible in the form

$$a_k \quad \text{or} \quad b_k = \frac{F(k)}{P_n(k)}, \qquad . \quad . \quad . \quad (7.10)$$

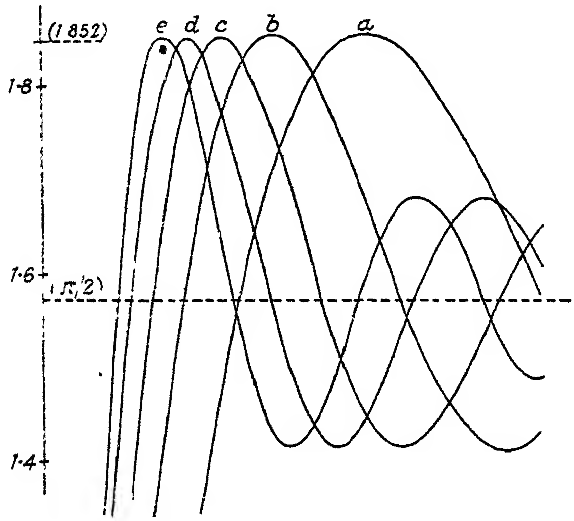


FIG. 12.—Part of the approximation curves for the function in Fig. 11*a*. (*a*)-(*e*) shows the sums of the first 5, 7, 9, 11 and 13 terms of the series (7.8), respectively.

where $F(k)$ is some function of the harmonic number k, frequently periodic (see, for example, formula (4.16), where

$$F = 16\ (1 - \cos k\pi)/\pi^3\ ),$$

and $P_n(k)$ is a polynomial function of degree n in k; in formula (4.16), $P_n(k) = k^3$ and $n = 3$. The rapidity of convergence of the series is controlled mainly by $P_n(k)$, and so depends largely upon the value of n. Now n has been shown to be equal to a certain reference number, determined thus: give to the original function represented by the series the reference number one, the first derivative the number two, the second derivative the number three, and so on; the numbering scheme is thus:

$$\begin{array}{ll} f(x) & 1, \\ \dfrac{d}{dx}f(x) & 2, \\ \dfrac{d^2}{dx^2}f(x) & 3, \\ \dfrac{d^3}{dx^3}f(x) & 4, \text{ etc.} \end{array}$$

Then n is the reference number of the lowest order derivative which is discontinuous.

This conclusion is admirably illustrated by the functions depicted in Figs. 5, 8 and 7, for which the Fourier coefficients are given by formulæ (4.4), (5.2) and (4.16). The square-peaked wave in Fig. 5 is itself discontinuous, and the value of n is therefore one; in (4.4) it is seen that in fact the coefficient a_k is, apart from the factor $\sin ka$, proportional to $1/k$, and the series is not very rapidly convergent. The saw-tooth wave in Fig. 8 is itself continuous, but its first derivative is discontinuous (there being violent changes in the slope of the graph at the angular points of the wave) so that $n = 2$; formula (5.2) shows that, apart from the factors representing $F(k)$ in (7.10), the values of a_k and b_k are proportional to $1/k^2$. Finally, the wave of parabolic arcs in Fig. 7 is continuous and has a continuous first derivative (the slope being such that there are no violent changes in direction along the graph) but the second derivative is discontinuous; formula (4.16) shows that, apart from the factor $1 - \cos k\pi$, the coefficient b_k is proportional to $1/k^3$ and the series is very rapidly convergent—so much so, in fact, that the parabolic-arc wave is a very good approximation to a sine-wave.

8. Other forms of Fourier series.

Two other forms of Fourier series are given brief mention below. They will be found useful in some engineering applications (see Karman and Biot, reference 11 in the Bibliography).

Complex form. Let

$$S = A_0 + \sum_{k=1}^{k=\infty} (a_k \cos kx + b_k \sin kx). \quad . \quad . \quad (8.1)$$

Now $\cos kx$ and $\sin kx$ can be written in the forms

$$\cos kx = \frac{e^{ikx} + e^{-ikx}}{2},$$

$$\sin kx = \frac{e^{ikx} - e^{-ikx}}{2i},$$

and hence (8.1) can be written as

$$\left.\begin{aligned} S &= \sum_{k=0}^{k=\infty} (C_k e^{ikx} + D_k e^{-ikx}) \\ \text{where} \quad C_0 &= A_0 \\ D_0 &= 0 \\ C_k &= \tfrac{1}{2}(a_k - ib_k) \\ D_k &= \tfrac{1}{2}(a_k + ib_k) \end{aligned}\right\} \quad . \quad . \quad . \quad (8.2)$$

or even more concisely as

$$\left.\begin{aligned} S &= \sum_{k=-\infty}^{k=\infty} C_k e^{ikx} \\ \text{where} \quad C_0 &= A_0 \\ C_k &= \tfrac{1}{2}(a_k - ib_k) \\ C_{-k} &= \tfrac{1}{2}(a_k + ib_k). \end{aligned}\right\} \quad . \quad . \quad . \quad (8.3)$$

Fourier integrals. It will be apparent, from a study of the foregoing sections of this chapter, that a Fourier series can be made to represent any periodic function which fulfils certain conditions of continuity; furthermore, it can be made to represent *any* function, not necessarily periodic, which fulfils these conditions of continuity, but if the function is not periodic the Fourier series will only approximate to it over the cycle of integration. It is a natural extension of the general theory to enquire whether this cycle of integration can be extended indefinitely, so that a function which is not periodic can be approximated by a Fourier series over an unrestricted range. The solution of this problem involves *Fourier's integrals.* The merest sketch of the derivation of these integrals is given below; for fuller details the reader is again referred to the works of Carslaw, and Karman and Biot (references 12, 13 in the Bibliography).

If a function is periodic with a period 2π radians it can easily be shown that the limits 0, 2π in the integrals connected with the Fourier coefficients (equations 3.3) can be replaced by $-\pi$, π, since that part of the cycle from $-\pi$ to 0 is identical with the part from π to 2π, and similarly for the functions $\cos kx$ and $\sin kx$. Furthermore, if the period of the function is $2l$, then if $f(x)$ is the function and

$$z = \frac{\pi x}{l},$$

it is seen that $f(x) = f\left(\frac{lz}{\pi}\right) = F(z)$ say, and

$$F(z) = A_0 + \sum_{k=1}^{k=\infty} (a_k \cos kz + b_k \sin kz), \quad . \quad . \quad (8.4)$$

where $$A_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(z)\,dz = \frac{1}{2l}\int_{-l}^{l} f(X)\,dX \quad . \quad . \quad (8.5a)$$

(the letter X is used in place of x to avoid confusion in the later stages; since the integral is definite its value is independent of the variable). Similarly,

$$a_l = \frac{1}{l}\int_{-l}^{l} f(X)\cos\frac{kX\pi}{l}\,dX, \quad . \quad . \quad (8.5b)$$

$$b_l = \frac{1}{l}\int_{-l}^{l} f(X)\sin\frac{kX\pi}{l}\,dX \quad . \quad . \quad (8.5c)$$

Substituting these values in (8.4),

$$f(x) = \frac{1}{2l}\int_{-l}^{l} f(X)\,.\,dX + \frac{1}{l}\sum_{k=1}^{k=\infty}\int_{-l}^{l} f(X)\left(\cos\frac{kX\pi}{l}\,.\,\cos\frac{kx\pi}{l} + \sin\frac{kX\pi}{l}\,.\,\sin\frac{kx\pi}{l}\right).\,dX$$

$$= \frac{1}{2l}\int_{-l}^{l} f(X)\,.\,dX + \frac{1}{l}\sum_{k=1}^{k=\infty}\int_{-l}^{l} f(X)\,.\,\cos\frac{k\pi}{l}(x - X)\,.\,dX. \quad (8.6)$$

If $\lambda_k = k\pi/l$, and $\lambda_k - \lambda_{k-1} = \pi/l = \Delta\lambda$, then the second term on the right-hand side of (8.6) is

$$\frac{1}{\pi}\sum_{k=1}^{k=\infty}\Delta\lambda\int_{-l}^{l} f(X)\,.\,\cos\lambda_l(x - X)\,.\,dX,$$

and if $\int_{-\infty}^{\infty} f(X)\,.\,dX$ is finite, then as $l \to \infty$, $\Delta\lambda \to 0$ and

$$f(x) = \frac{1}{\pi}\int_0^{\infty} d\lambda \int_{-\infty}^{\infty} f(X)\,.\,\cos\lambda\,(x - X)\,.\,dX. \quad . \quad (8.7)$$

The expression in (8.7) is termed a *Fourier's integral*. These integrals are useful in the analysis of the elastic properties of beams of very great span, to mention only one practical application.

9. Two integration results.

$$\text{(i)} \quad \int_0^x \frac{\sin 2nt}{t}\,dt = \int_0^{2nx} \frac{\sin t}{t}\,dt. \quad . \quad . \quad (9.1)$$

This result, utilised on page 176 above, is easily proved.

Let $y = 2nt$, then $dt = dy/2n$; and when $t = x$, $y = 2nx$. Hence

$$\int_0^x \frac{\sin 2nt}{t}\,dt = \int_0^{2nx} \frac{\sin y}{y}\,dy$$
$$= \int_0^{2nx} \frac{\sin t}{t}\,dt,$$

since the definite integral is unaltered in value no matter what letter is used to represent the variable.

$$\text{(ii)} \int_0^\infty \frac{\sin t}{t}\, dt = \pi/2. \quad . \quad . \quad . \quad (9.2)$$

The formal proof of this result requires a thorough knowledge of the continuity of functions and validity of integrals; for these matters the reader is referred to any textbook on the theory of functions. In this section it will be assumed that all integrations performed are valid.

Since $\underset{a\to 0}{\text{Lt.}}\ \epsilon^{-at} = 1$,

$$\int_0^\infty \frac{\sin t}{t}\, dt = \underset{a\to 0}{\text{Lt.}} \int_0^\infty e^{-at} \frac{\sin t}{t}\, dt,$$

and if
$$F(a) = \int_0^\infty e^{-at} \frac{\sin t}{t}\, dt,$$

$$F'(a) = \frac{\partial}{\partial a} F(a) = \int_0^\infty \frac{\partial}{\partial a}\left(e^{-at}\frac{\sin t}{t}\right) dt$$

$$= -\int_0^\infty e^{-at} \sin t\, .\, dt.$$

Now,
$$\frac{d}{dt}\epsilon^{-at}(\cos t + a \sin t) = -(a^2+1)e^{-at}\sin t,$$

and hence
$$\int_0^\infty e^{-at} \sin t\, .\, dt = -\frac{1}{a^2+1}\Big[e^{-at}(\cos t + a \sin t)\Big]_0^\infty$$

$$= \frac{1}{a^2+1}.$$

Thus
$$F'(a) = -\frac{1}{a^2+1},$$

and
$$F(a) = \pi/2 - \tan^{-1} a$$

(since
$$\underset{a\to\infty}{\text{Lt.}}\ F(a) = 0).$$

Thus
$$\int_0^\infty \frac{\sin t}{t}\, dt = \underset{a\to 0}{\text{Lt.}}\ (\pi/2 - \tan^{-1} a),$$

and the result (9.2) follows.

CHAPTER VII

NUMERICAL METHODS

1. Introductory.

The mathematical method of analysis described in Chapter VI is limited, in its immediate application, to that type of waveform which can be regarded as composed of segments of simple mathematical curves—straight line segments, parabolic arcs, and arcs of sine-waves, for example—so that the necessary integrations can be performed exactly. Some of the periodic variations occurring in engineering problems are of this type; for example, the force transmitted during a hammer-blow applied to a structure can be represented approximately by a cycle of the saw-tooth wave analysed on page 160, and the effects of a regular repetition of such blows, including the so-called "recurring transients," can be determined by straightforward mathematical methods (see reference 1 in the Bibliography at the end of the book). Similarly, the tangential effort variation in a single cylinder of an internal-combustion engine can be represented very roughly by an intermittent sine-wave of the type analysed on page 163.

The majority of waveforms and periodic variations encountered in engineering and scientific research, however, are not amenable to this treatment, and in such cases recourse must be had, in general, to the numerical method described in this chapter. There are two distinct ways of regarding the numerical method: first, as an approximation to the integration method employing the Cauchy integrals, and secondly, as a process of curve-fitting. The method itself, and the results obtained by its use, are of course unaffected by the choice of point of view; but it is always advisable to keep the dual nature of the process in mind when applying or interpreting the results.

A periodic function can be expressed as a Fourier series in only one way, the coefficients being determined by the Cauchy integrals. The fact that the Fourier series representation of a periodic function is unique is known as Riemann's Theorem, an informal demonstration of which is given in Appendix II (p. 247); the rigorous proof is given by Whittaker and Watson (reference 2 in the Bibliography at the end of the book). The theorem applies only to a series in which no limit is imposed upon the number of terms which may be included; if any arbitrary limitation is imposed, the result of

the theorem is not necessarily true. Such a limitation is involved in the numerical method of analysis, as may best be seen by regarding the process as one of curve-fitting. Suppose that 24 ordinates are drawn to the curve representing the periodic variation, the ordinates being conveniently distributed evenly over the cycle. The points at which these ordinates intersect the curve will be termed hereafter the "selected points." Now, an unlimited number of different curves can be drawn to pass through the 24 selected points; in particular, one curve will represent the sum of a harmonic series which contains a constant term, the first eleven sine components and the first twelve cosine components. It is this series which is determined by the normal 24-ordinate scheme of analysis. It will at once be evident that if the unique *unlimited* Fourier series expansion of the variation contains harmonic components of higher order than the 11th (sine) and 12th (cosine), the coefficients for these higher harmonics will not be determined by such a process. What is not so evident, and has been the cause of much misapprehension concerning the accuracy of numerical harmonic analyses, is that in such a case the coefficients of the lower harmonics, as determined by the analysis, may be grossly in error.

In several textbooks on engineering mathematics, and in other memoirs, there appear statements which imply that the number of ordinates used in a numerical analysis must exceed twice the reference number of the highest harmonic whose coefficients it is required to determine; such statements err, not by reason of complete falsity, but by suppression of the most important aspect of the matter

in actual fact, the number of ordinates must exceed twice the reference number of the highest harmonic present in the unlimited and unique Fourier series expansion of the variation, and this number may be far greater than the one quoted above.

The curve representing the sum of the limited series found by the numerical method will certainly pass through the 24 selected points (or however many points were selected for the analysis), but any correspondence between the sum of this series and the original variation at other points in the cycle is purely coincidental if the unlimited Fourier series contains harmonics of higher order than those specified above, and the analysis will not yield an accurate assessment of the harmonic contents of the variation.

If the function actually contains components of, say, the 1st,

3rd, 7th, 10th and 11th, harmonics, the coefficients of these terms will be given accurately by a 24-ordinate scheme of analysis. If, however, there are present in the function components of, say, the 13th, 14th and 20th harmonics as well, the series found will contain components of the 1st, 3rd, 4th, 7th, 10th and 11th harmonics, and of these the coefficients of the 1st, 3rd and 7th alone will be correct. With a 24-ordinate scheme of analysis, the 13th, 14th and 20th harmonics are confused with the 11th, 10th and 4th harmonics respectively, and on the basis of such an analysis alone it is not possible to determine whether an indicated 4th harmonic, for example, really represents a 4th or a 20th (or, for that matter, a 28th, 44th, 52nd or 68th, etc.) harmonic, or some combination of these.

Naturally, the extent to which results are rendered invalid by this phenomenon depends upon the amplitudes of the higher harmonics in the true Fourier series, and several examples have been included in this chapter to illustrate typical cases. Unfortunately, several engineering textbooks in describing the method have completely disregarded this matter, and in some cases have actually given misleading statements concerning the accuracy of numerical analyses. For this reason, a detailed discussion of the matter has been given in Sections 4 and 5, pages 189, 192.

A 48-ordinate scheme of analysis, which is capable of giving accurate results for functions of which the harmonics are not higher than the 23rd (sine) and 24th (cosine) components, is sufficient for most practical purposes, since harmonics higher than these (if they are present at all) are usually so small as to produce negligible errors. A 72-ordinate scheme, ascribed to Denman and Withers, is given by Stansfield (reference 3 in the Bibliography).

The development of systematic schemes of numerical analysis seems to have originated with Runge (reference 4). Many forms have been devised for various special problems, and the reader will find detailed references to the more important descriptions of the methods in the Handbook of the Napier Tercentenary Exhibition (reference 5).

The chapter deals exclusively with methods of numerical analysis utilising a number of ordinates spaced evenly over the cycle. Other special methods have been developed, wherein the ordinates considered are spaced at uneven intervals, according to a special pattern; for descriptions of these methods the reader should consult references 6 and 7 in the Bibliography at the end of the book.

2. Derivation of formulæ from mathematical method.

The numerical method of harmonic analysis is used when the periodic function to be analysed is given in such a form that its value is known at various points in the cycle. These values may be the direct results of physical observation or may be interpolated graphically from such observations. For the sake of simplicity in numerical work it is usual to take as data the values of the function at an even number of points evenly distributed over the cycle.

Let $f(x)$ be a continuous function to be analysed; if 2π radians is the period of $f(x)$, then $f(x-2\pi)=f(x)$. Let ordinates to the curve $y=f(x)$ be drawn at the points $x_0=0$, x_1, x_2, . . . x_{N-1}, $x_N=2\pi$, such that

$$x_r = r\,.\,\Delta x, \quad \Delta x = 2\pi/N, \quad . \quad . \quad . \quad (2.1)$$

and let the corresponding values of the ordinates be y_0, y_1, y_2, . . . y_N. Since $f(x)$ is a continuous function of x with period 2π radians,

$$y_0 = y_N.$$

The formulæ (3.3) of Chapter VI (p. 143) can be regarded as limiting expressions in the following manner:

$$\frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\,.\,dx = \underset{N\to\infty}{\mathrm{Lt.}}\ \frac{1}{\pi}\sum_{r=1}^{r=N} y_r \cos kx_r\,.\,\Delta x$$

$$= \underset{N\to\infty}{\mathrm{Lt.}}\ \frac{2}{N}\sum_{r=1}^{r=N} y_r \cos kx_r.$$

Hence the cosine coefficient is given approximately by the formula

$$\left.\begin{aligned} a_k &= \frac{2}{N}\sum_{r=1}^{r=N} y_r \cos kx_r \\ \text{and similarly,}\quad b_k &= \frac{2}{N}\sum_{r=1}^{r=N} y_r \sin kx_r \end{aligned}\right\} \quad . \quad . \quad . \quad (2.2)$$

This derivation is quite informal and requires to be substantiated by further reasoning; in particular a lower limit must be set to the value of N (i.e. the number of ordinates taken) in accordance with the number of harmonics present in the function $f(x)$. These matters are considered in Sections 4 and 5 below. The procedure summarised in the formulæ (2.2) can be stated thus:

"The coefficient a_k in the series

$$S = A_0 + \sum_{k=1}^{k=n}(a_k \cos kx + b_k \sin kx), \quad . \quad . \quad (2.3)$$

representing the function $y=f(x)$ is found by dividing the cycle into N equal divisions, measuring the ordinates of the curve $y=f(x)$ at these N points, multiplying each ordinate y_r by the corresponding

value $\cos kx_r$ of the function $\cos kx$, summing the products so formed for all the ordinates and multiplying the sum by $2/N$.

"The coefficient b_k is found in a similar manner, the function $\sin kx$ being substituted for $\cos kx$.

"The constant term $A_0 = a_0/2$, where a_0 is found as for a_k, putting $k = 0$; i.e. A_0 is the average value of the ordinates."

It should be noted that, as shown on page 193, if N is even the value given by (2.2) for a_k, when $k = N/2$, is twice the correct value.

3. Simple example.

Expressing the periodic variation in the form (2.3) slightly modified, i.e.

$$y = f(x) = A_0 + \sum_k a_k \cos kx + \sum_k b_k \sin kx, \qquad (3.1)$$

where the period is 2π radians or 360° in x, suppose that the values of y for $x = 0$, $\pi/2$ π, and $3\pi/2$ are known or measured from a record of the waveform.

Denoting these values by y_0, y_1, y_2, y_3 respectively,

$$\left.\begin{aligned} y_0 &= f(0) = A_0 + \sum_k a_k \\ y_1 &= f(\pi/2) = A_0 + \sum_k a_k \cos k\pi/2 + \sum_k b_k \sin k\pi/2 \\ y_2 &= f(\pi) = A_0 + \sum_k a_k \cos k\pi \\ y_3 &= f(3\pi/2) = A_0 + \sum_k a_k \cos 3k\pi/2 + \sum_k b_k \sin 3k\pi/2 \end{aligned}\right\} \qquad (3.2)$$

Now consider the effect of summing these four values. Denoting the sum by c_0,

$$\left.\begin{aligned} c_0 &= y_0 + y_1 + y_2 + y_3 = 4A_0 + \Sigma a_k G_k + \Sigma b_k H_k \\ \text{where } G_k &= 1 + \cos k\pi/2 + \cos k\pi + \cos 3k\pi/2 \\ H_k &= \sin k\pi/2 + \sin 3k\pi/2 \end{aligned}\right\} \qquad (3.3)$$

The value of c_0 clearly depends upon the value of G_k and H_k. Now if k is a multiple of 4, say 4P where P is an integer, then

$$\begin{aligned} \cos k\pi/2 &= \cos 2P\pi = 1, \\ \cos k\pi &= \cos 4P\pi = 1, \\ \cos 3k\pi/2 &= \cos 6P\pi = 1, \end{aligned}$$

so that $G_{4P} = 4$; similarly, it can be shown that $H_{4P} = 0$.

If $k = 4P - 1$, it is found that $G_{4P-1} = 0$, and $H_{4P-1} = 0$, and similarly, $G_{4P-2} = H_{4P-2} = G_{4P-3} = H_{4P-3} = 0$. Substituting these values in (3.3),

$$c_0 = 4(A_0 + \sum_P a_{4P}). \qquad (3.4)$$

Again, consider the effect of performing the summation

$$\begin{aligned} c_1 &= y_0 \cos 0 + y_1 \cos \pi/2 + y_2 \cos \pi + y_3 \cos 3\pi/2 \\ &= y_0 - y_2 \\ &= \sum_k a_k (1 - \cos k\pi). \end{aligned}$$

If k is even, $\cos k\pi = 1$ and $(1 - \cos k\pi) = 0$; but if k is odd, $\cos k\pi = -1$ and $(1 - \cos k\pi) = 2$. Representing odd values of k by $2P - 1$,

$$c_1 = 2 \sum_P a_{2P-1}. \qquad (3.5)$$

Similarly, the summation

$$\begin{aligned} s_1 &= y_0 \sin 0 + y_1 \sin \pi/2 + y_2 \sin \pi + y_3 \sin 3\pi/2 \\ &= y_1 - y_3, \end{aligned}$$

yields

$$\left.\begin{aligned} s_1 &= \Sigma a_k G'_k + \Sigma b_k H'_k \\ \text{where} \quad G'_k &= \cos k\pi/2 - \cos 3k\pi/2 \\ H'_k &= \sin k\pi/2 - \sin 3k\pi/2 \end{aligned}\right\} \qquad (3.6)$$

$$\begin{aligned} \text{If } k = 4P, \quad & G'_{4P} = \cos 2\pi - \cos 6\pi = 0, \\ & H'_{4P} = \sin 2\pi - \sin 6\pi = 0. \\ \text{If } k = 4P - 1, \quad & G'_{4P-1} = \cos \pi/2 - \cos 3\pi/2 = 0, \\ & H'_{4P-1} = -\sin \pi/2 + \sin 3\pi/2 = -2. \\ \text{If } k = 4P - 2, \quad & G'_{4P-2} = \cos \pi - \cos 3\pi = 0, \\ & H'_{4P-2} = -\sin \pi + \sin 3\pi = 0. \\ \text{If } k = 4P - 3, \quad & G'_{4P-3} = \cos 3\pi/2 - \cos \pi/2 = 0, \\ & H'_{4P-3} = -\sin 3\pi/2 - \sin \pi/2 = 2. \end{aligned}$$

Substituting these results in (3.6),

$$s_1 = 2 \sum_P (b_{4P-3} - b_{4P-1}). \qquad (3.7)$$

Proceeding in a similar manner to perform the summations

$$\begin{aligned} c_2 &= y_0 \cos 0 + y_1 \cos \pi + y_2 \cos 2\pi + y_3 \cos 3\pi \\ &= y_0 - y_1 + y_2 - y_3, \end{aligned}$$

and

$$\begin{aligned} s_2 &= y_0 \sin 0 + y_1 \sin \pi + y_2 \sin 2\pi + y_3 \sin 3\pi \\ &= 0, \end{aligned}$$

the results are obtained :

$$\left.\begin{aligned} c_2 &= 4 \sum_P a_{4P-2}, \\ s_2 &= 0. \end{aligned}\right\} \qquad (3.8)$$

The results (3.4, 5, 7, 8) are expanded below :

$$\left.\begin{aligned} c_0 &= (4A_0) + 4a_4 + 4a_8 + \text{etc.} \\ c_1 &= (2a_1) + 2a_3 + 2a_5 + 2a_7 + \text{etc.} \\ s_1 &= (2b_1) - 2b_3 + 2b_5 - 2b_7 + \text{etc.} \\ c_2 &= (4a_2) + 4a_6 + 4a_{10} + \text{etc.} \end{aligned}\right\} \quad . \quad . \quad (3.9)$$

The most important point to be noted is that if the variation (3.1) contains harmonics higher than the 2nd, no information concerning the coefficients of *any* harmonics is afforded by the equations (3.9) ; if, however, the variation consists merely of the constant term A_0 and the first two harmonics, i.e.

$$y = A_0 + a_1 \cos x + a_2 \cos 2x + b_1 \sin x + b_2 \sin 2x, \quad (3.10)$$

then the terms outside the brackets on the right-hand side of the equations (3.9) are zero, and these equations then yield the results :

$$\left.\begin{aligned} A_0 &= c_0/4 \\ a_1 &= c_1/2 \\ a_2 &= c_2/4 \\ b_1 &= s_1/2 \end{aligned}\right\} \quad . \quad . \quad . \quad (3.11)$$

No information has been obtained concerning the value of the coefficient b_2 of the second sine component, since this component has zero value for $x = 0$, $\pi/2$, π, and $3\pi/2$.

4. False indications from using too few ordinates.

The processes by which the quantities c_0, etc., were found in the preceding section are approximations to the Cauchy integrations

$$\pi \,.\, a_k = \int_0^{2\pi} f(x) \cos kx \,.\, dx,$$

$$A_0 = a_0/2,$$

$$\pi \,.\, b_k = \int_0^{2\pi} f(x) \sin kx \,.\, dx,$$

except for the numerical constant by which the coefficient a_2 is multiplied (see Section 10). Whereas these integrations yield accurate results, however, the numerical approximations do not, unless the 3rd and higher harmonics are absent. In general, as will be shown in the following section, accurate results will only be obtained if the number of ordinates used is greater than twice the reference number of the highest harmonic present in the waveform or variation. It is quite true that the function

$$y' = c_0/4 + (c_1/2) \cos x + (c_2/4) \cos 2x + (s_1/2) \sin x, \quad (4.1)$$

found by substituting the results (3.11) in the equation (3.10)

assumes the correct values y_0, y_1, y_2, and y_3 at the values $x = 0$, $\pi/2$, π, $3\pi/2$, regardless of the presence or absence of higher harmonics: for example, at $x = \pi/2$,

$$\begin{aligned} y'(\pi/2) &= c_0/4 - c_2/4 + s_1/2 \\ &= (y_0 + y_1 + y_2 + y_3)/4 - (y_0 - y_1 + y_2 - y_3)/4 + (y_1 - y_3)/2 \\ &= y_1. \end{aligned}$$

If, however, the 3rd or higher harmonics are present, the function (4.1) does not assume the correct values at any other points in the cycle; or at any rate, any correct value assumed for values of x other than 0, $\pi/2$, π or $3\pi/2$ is purely coincidental and not inherent in the theory of the method.

The practical importance of this result cannot be over-emphasised. In several reputable engineering textbooks a statement may be found, which expresses in effect the opinion that if the coefficients of only a few of the lower harmonics are to be determined, a simple method of numerical analysis can successfully be employed, where the number of ordinates used is just greater than twice the reference number of the highest harmonic whose coefficients are required; for example, it is sometimes contended that if only the first ten harmonics are required a scheme of analysis using 24 ordinates will give accurate results. *Such a statement is definitely erroneous and misleading.* If it were correct the process outlined in the preceding section, which utilises four ordinates, would on all occasions yield the correct values for A_0, a_1, and b_1, whereas the form of the equations (3.9) shows that these values will only be obtained if the harmonics higher than the second are absent.

Consider the variation illustrated in Fig. 9*b*, Chapter VI (p. 164). The four ordinates used in the method of the preceding section are $y_0 = 1$, $y_1 = 0$, $y_2 = 0$, $y_3 = 0$. Thus $c_0 = 1$, $c_1 = 1$, $c_2 = 1$, $s_1 = 0$. The false indications obtained by assuming these quantities to be respectively $4A_0$, $2a_1$, $4a_2$, and $2b_1$ would be

$$A_0 = 0{\cdot}25, \quad a_1 = 0{\cdot}5, \quad a_2 = 0{\cdot}25, \quad b_1 = 0.$$

The correct values are given in Table III, page 166, as

$$A_0 = 0{\cdot}319, \quad a_1 = 0{\cdot}5, \quad a_2 = 0{\cdot}212, \quad b_1 = 0.$$

Utilising the other values given in the table, it is seen that

$$\begin{aligned} &4A_0 + 4a_4 + 4a_8 + \text{etc.} \\ &\qquad = 1{\cdot}274 - 0{\cdot}168 - 0{\cdot}040 - 0{\cdot}016 - \text{etc.} \\ &\qquad = 1{\cdot}050 - \text{some small quantity.} \\ &2a_1 + 2a_3 + 2a_5 + \text{etc.} = 1{\cdot}000. \\ &4a_2 + 4a_6 + 4a_{10} + \text{etc.} \\ &\qquad = 0{\cdot}848 + 0{\cdot}072 + 0{\cdot}024 + \text{etc.} \\ &\qquad = 0{\cdot}944 + \text{some small quantity.} \\ &2b_1 - 2b_3 + 2b_5 - \text{etc.} = 0. \end{aligned}$$

The quantities 1·050 — and 0·944 + both become 1·000 when the harmonics higher than the 12th (not given in the table) are taken into account, and thus the equations (3.9) are verified for this particular example. The errors involved by the erroneous supposition that only the terms inside the brackets on the right-hand side of these equations are represented by the quantities on the left-hand side are indicated below:

A_0: true value 0·319, false value 0·25, error 22 per cent.
a_2: true value 0·212, false value 0·25, error 18 per cent.
a_1 and b_1 are correct.

The absence of error in the values for a_1 and b_1 is due to the complete absence of the sine components, since the wave is symmetrical, and the absence of all odd cosine components other than that of the 1st harmonic.

Again, consider the variation illustrated in Fig. 9*c* on page 164. In this case the four ordinates are $y_0 = 1$, $y_1 = 0{\cdot}707$, $y_2 = 0$, $y_3 = 0{\cdot}707$. Neglecting the sine components, which are all zero since the wave is symmetrical, the values of c_0, etc., are

$$c_0 = 2{\cdot}414, \quad c_1 = 1, \quad c_2 = -\,0{\cdot}414,$$

and if these are supposed to be the values of $4A_0$, $2a_1$, and $4a_2$, respectively, the false results are:

$$A_0 = 0{\cdot}604, \quad a_1 = 0{\cdot}500, \quad a_2 = -\,0{\cdot}104.$$

The true results are given in the last column of Table III, page 166, and it is seen that the errors are as indicated below:

A_0: true value 0·637, false value 0·604, error 5 per cent.
a_1: true value 0·425, false value 0·500, error 15 per cent.
a_2: true value — 0·085, false value — 0·104, error 22 per cent.

Moreover,

$$\begin{aligned}
&4A_0 + 4a_4 + 4a_8 + \text{etc.}\\
&\qquad = 2{\cdot}548 - 0{\cdot}080 - 0{\cdot}020 - 0{\cdot}008 - \text{etc.}\\
&\qquad = 2{\cdot}440 - \text{a small quantity.}\\
&2a_1 + 2a_3 + 2a_5 + \text{etc.}\\
&\qquad = 0{\cdot}850 + 0{\cdot}072 + 0{\cdot}026 + 0{\cdot}014 + 0{\cdot}008\\
&\qquad\qquad + 0{\cdot}006 + \text{etc.}\\
&\qquad = 0{\cdot}976 + \text{a small quantity.}\\
&4a_2 + 4a_6 + 4a_{10} + \text{etc.}\\
&\qquad = -\,0{\cdot}340 - 0{\cdot}036 - 0{\cdot}012 - \text{etc.}\\
&\qquad = -\,0{\cdot}388 - \text{a small quantity.}
\end{aligned}$$

The values 2·440 —, 0·976 +, and — 0·388 — become 2·414, 1, — 0·414 (i.e. c_0, c_1 and c_2) when the harmonics higher than the 12th are included.

5. Alternative derivation from process of curve-fitting.

The series (2.3) may be regarded as an approximation to the function $y = f(x)$ and the coefficients a_k and b_k determined by consideration of *error*. Putting $x = x_1, x_2 \ldots x_N$ in turn in (2.3) gives the equations

$$\left.\begin{aligned} S_1 &= A_0 + \sum_{k=1}^{k=n} (a_k \cos kx_1 + b_k \sin kx_1) \\ S_2 &= A_0 + \sum_{k=1}^{k=n} (a_k \cos kx_2 + b_k \sin kx_2) \\ &\text{etc., etc.,} \\ S_N &= A_0 + \sum_{k=1}^{k=n} a_k \end{aligned}\right\} \quad . \quad . \quad (5.1)$$

and from these N equations it is required to determine the values of the $(2n + 1)$ coefficients $A_0, a_1 \ldots a_n, b_1 \ldots b_n$. It follows that N must not be less than $(2n + 1)$, since as many equations are required as there are unknown quantities to be determined. If $N = 2n + 1$, the coefficients are uniquely determined by the equations (5.1) and it is easy to show that their values are given by (2.2). If $N > 2n + 1$ it may not be possible to solve the equations (5 1) exactly; for example, it is not possible to find quantities p and q satisfying the three equations

$$\begin{aligned} p - q &= 1 \qquad (a), \\ 2p - q &= 4 \qquad (b), \end{aligned}$$

and

$$p - 2q = -4 \quad (c).$$

Taking equations (a) and (b), the values $p = 3$, $q = 2$ are found; taking (a) and (c), $p = 6$. $q = 5$; and taking (b) and (c), $p = q = 4$. The three equations between two unknowns are in this case inconsistent, as may be seen at once from the fact that by adding (b) and (c) together the result $p - q = 0$ is obtained, and this is inconsistent with (a). On the other hand, if the third equation had been $p + q = 5$ the solution $p = 3$, $q = 2$ would have satisfied all three equations.

When $N > 2n + 1$, i.e. when the number of ordinates exceeds one more than twice the reference number of the highest harmonic present, or supposed to be present, in the variation, recourse is had

to the principle of *least mean squared error*, to determine the best values for the coefficients so that the curve $y = f(x)$ is most nearly approximated by the curve $y' = S(x)$. Let S_r be the value of the series S for $x = x_r$, $(r = 1, 2 \ldots N)$, then $(y_r - S_r)$ is the difference between the value of the function $f(x)$ and the series $S(x)$ for $x = x_r$,

and $$E = \frac{1}{N} \sum_{r=1}^{N}{}' (y_r - S_r)^2 \quad . \quad . \quad . \quad (5.2)$$

is the mean squared error involved in using the series $S(x)$ in place of the function $f(x)$. The coefficients are found by solving the equations

$$\left.\begin{aligned} \frac{\partial E}{\partial a_k} &= 0 \\ \frac{\partial E}{\partial b_k} &= 0 \end{aligned}\right\}, \quad . \quad . \quad . \quad . \quad . \quad (5.3)$$

which give a maximum or minimum value to E, in a manner similar to that used in Section 7 of Chapter VI (p. 173), and it is found that the solutions give the formulæ (2.2), with a trivial exception: if N is even, the value of a_k where $k = N/2$ can be found, and is just half the value given by the formula (2.2). In this process the following equations take the place of equations (3.5) of Chapter VI (p. 144):

$$\left.\begin{aligned} &\sum_{r=1}^{r=N} \sin mx_r \, . \sin nx_r = 0, \text{ etc., etc.} \\ &\text{if } m \text{ and } n \text{ are unequal integers, and } m + n < N, \\ &\sum_{r=1}^{r=N} \sin kx_r \, . \cos kx_r = 0 \\ &\sum_{r=1}^{r=N} \sin^2 kx_r = N/2 \text{ if } k \neq N/2 \\ &\qquad\qquad = 0 \quad \text{ if } k = N/2 \\ &\sum_{r=1}^{r=N} \cos^2 kx_r = N/2 \text{ if } k \neq N/2 \\ &\qquad\qquad = N \quad \text{ if } k = N/2 \\ &\text{if } k \text{ is an integer.} \end{aligned}\right\} \quad . \quad . \quad . \quad (5.4)$$

it being understood that x_r fulfils the condition (2.1) (p. 186). The equations (5.4) are proved in Section 10, below (p. 210). The argument given in Chapter VI to show that the error is a minimum rather than a maximum is repeated here for convenience: since

the coefficients determined are all finite if the function $f(x)$ is finite throughout the cycle, as it is in all practical cases, a worse approximation can be found by making all the coefficients very large; the conditions (5.3) therefore give a minimum value to E.

It is now possible to examine the general solution to determine why it is essential to use N ordinates when the variation includes harmonics up to and including the nth, where N is greater than $2n$. Suppose that it is required merely to determine the coefficients of the harmonics up to and including the mth, where $m < n$. Utilising for this purpose M ordinates, where M is greater than $2m$, M equations corresponding to (5.1) are obtained, and these equations contain $(2n + 1)$ unknown quantities, A_0, $a_1 \ldots a_n$, $b_1 \ldots b_n$. It will not be possible to solve the equations completely unless there are as many equations as there are unknown quantities, i.e. unless M is greater than $2n$. Owing to the peculiar nature of the equations, it will in all cases be possible to solve them in terms of $(2m + 1)$ quantities c_0', $c_1' \ldots c_m'$, $s_1' \ldots s_m'$, where

$$\left.\begin{aligned}
c_0' &= A_0 + a_M + a_{2M} + \text{etc.} \\
c_1' &= a_1 + a_{M-1} + a_{M+1} + a_{2M-1} + a_{2M+1} + \text{etc.} \\
c_2' &= a_2 + a_{M-2} + a_{M+2} + a_{2M-2} + a_{2M+2} + \text{etc.,} \\
&\quad \text{etc.} \\
s_1' &= b_1 - b_{M-1} + b_{M+1} - b_{2M-1} + b_{2M+1} - \text{etc.} \\
s_2' &= b_2 - b_{M-2} + b_{M+2} - b_{2M-2} + b_{2M+2} - \text{etc.,} \\
&\quad \text{etc.}
\end{aligned}\right\} \qquad (5.5)$$

In the critical case $M = 2m + 1$, this solution can be effected by ordinary elimination: when M exceeds $2m + 1$, that is when there are more equations than there are unknown quantities, the method of least squared error is used. It will be found that for $M > 2n + 1$, i.e. $m > n$, the quantities (5.5) degenerate to the coefficients A_0, $a_1 \ldots a_n$, $b_1 \ldots b_n$.

6. Simple example of application of method: tabular scheme.

Consider the function

$$y = A_0 + a_1 \cos x + a_2 \cos 2x + b_1 \sin x + b_2 \sin 2x. \qquad (6.1)$$

For the complete analysis of a given function of this type it is necessary to take at least five ordinates, since the 2nd harmonic may be present. For convenience, six ordinates are taken, at the points $x = 0$, 60, 120, 180, 240 and 300 degrees. The corresponding values of $\cos kx_r$ and $\sin kx_r$ are given in Table I for $r = 0, 1 \ldots 5$ and $k = 1, 2$.

TABLE I

Values of $\binom{\cos}{\sin} x_r$ and $\binom{\cos}{\sin} 2x_r$ for $x_r = r \cdot 60^\circ$

x_r	$\sin x_r$	$\cos x_r$	$\sin 2x_r$	$\cos 2x_r$
0	0	1	0	1
60°	0·866	0·500	0·866	−0·500
120°	0·866	−0·500	−0·866	−0·500
180°	0	−1	0	−1
240°	−0·866	−0·500	0·866	−0·500
300°	−0·866	0·500	−0·866	−0·500

Let the corresponding values of y be $y_0 \ldots y_5$. Then application of the formulæ (2.2) gives the results :

$$\left.\begin{aligned}
A_0 &= \tfrac{1}{6}(y_0 + y_1 \ldots + y_5).\\
a_1 &= \tfrac{1}{3}(y_0 \cos 0 + y_1 \cos 60^\circ \ldots + y_5 \cos 300^\circ)\\
&= \tfrac{1}{3}[0{\cdot}500\,(y_1 - y_2 - y_4 + y_5) + y_0 - y_3].\\
a_2 &= \tfrac{1}{3}(y_0 \cos 0 + y_1 \cos 120^\circ \ldots + y_5 \cos 600^\circ).\\
&= \tfrac{1}{3}[-0{\cdot}500(y_1 + y_2 + y_4 + y_5) + y_0 + y_3].\\
b_1 &= \tfrac{1}{3}(y_0 \sin 0 + y_1 \sin 60^\circ \ldots + y_5 \sin 300^\circ)\\
&= \tfrac{1}{3}[0{\cdot}866(y_1 + y_2 - y_4 - y_5)].\\
b_2 &= \tfrac{1}{3}(y_0 \sin 0 + y_1 \sin 120^\circ \ldots + y_5 \sin 600^\circ)\\
&= \tfrac{1}{3}[0{\cdot}866(y_1 - y_2 + y_4 - y_5)].
\end{aligned}\right\} \qquad (6.2)$$

These equations enable the five coefficients to be found.

In the particular case where the ordinates are

y_0	y_1	y_2	y_3	y_4	y_5
5·0	11·3	− 0·4	− 1·0	1·4	− 4·3

the coefficients are given by (6.2) as :

$$\begin{aligned}
A_0 &= \tfrac{1}{6}(5{\cdot}0 + 11{\cdot}3 - 0{\cdot}4 - 1{\cdot}0 + 1{\cdot}4 - 4{\cdot}3) = 2{\cdot}0,\\
a_1 &= \tfrac{1}{3}[0{\cdot}500\,(11{\cdot}3 + 0{\cdot}4 - 1{\cdot}4 - 4{\cdot}3) + 5{\cdot}0 + 1{\cdot}0] = 3{\cdot}0,\\
a_2 &= \tfrac{1}{3}[-\,0{\cdot}500\,(11{\cdot}3 - 0{\cdot}4 + 1{\cdot}4 - 4{\cdot}3) + 5{\cdot}0 - 1{\cdot}0] = 0,\\
b_1 &= \tfrac{1}{3}[0{\cdot}866\,(11{\cdot}3 - 0{\cdot}4 - 1{\cdot}4 + 4{\cdot}3)] = 4{\cdot}0\\
b_2 &= \tfrac{1}{3}[0{\cdot}866\,(11{\cdot}3 + 0{\cdot}4 + 1{\cdot}4 + 4{\cdot}3)] = 5{\cdot}0
\end{aligned}$$

The series representing the function is therefore

$$y = 2 + 3 \cos x + 4 \sin x + 5 \sin 2x.$$

Tabular scheme. The simple example given above can be treated quite adequately by the method described, but tabular schemes have been evolved for dealing with more complicated functions (in the analysis of which a larger number of ordinates are taken), and the development of these tabular schemes can be illustrated

by applying the process first to this simple case where only six ordinates are taken.

The ordinates are written down as shown below, and the sums V and differences W formed, so that $V_1 = y_1 + y_5$, and $W_1 = y_1 - y_5$, etc.:

$$
\begin{array}{lcccc}
 & y_0 & y_1 & y_2 & y_3 \\
 & & y_5 & y_4 & \\
\hline
\text{sums} & V_0 & V_1 & V_2 & V_3 \\
\text{differences} & - & W_1 & W_2 & -
\end{array}
$$

From the V's and W's the sums and differences P, Q, R, S and L are formed:

$$
\begin{array}{lcc|cc}
 & V_0 & V_1 & W_1 & P_0 \\
 & V_3 & V_2 & W_2 & P_1 \\
\hline
\text{sums} & P_0 & P_1 & R & L \\
\text{differences} & Q_0 & Q_1 & S & -
\end{array}
$$

Then the coefficients are given by the formulæ:

$$
\left.
\begin{aligned}
6A_0 &= L \\
3a_1 &= 0{\cdot}500\ Q_1 + Q_0 \\
3a_2 &= -\ 0{\cdot}500\ P_1 + P_0 \\
3b_1 &= 0{\cdot}866\ R \\
3b_2 &= 0{\cdot}866\ S
\end{aligned}
\right\} \quad . \quad . \quad . \quad (6.3)
$$

For

$$
\begin{aligned}
P_0 &= V_0 + V_3 = y_0 + y_3, \\
P_1 &= V_1 + V_2 = y_1 + y_2 + y_4 + y_5, \\
Q_0 &= V_0 - V_3 = y_0 - y_3, \\
Q_1 &= V_1 - V_2 = y_1 - y_2 - y_4 + y_5, \\
R &= W_1 + W_2 = y_1 + y_2 - y_4 - y_5, \\
S &= W_1 - W_2 = y_1 - y_2 + y_4 - y_5, \\
L &= P_0 + P_1 = y_0 + y_1 + y_2 + y_3 + y_4 + y_5.
\end{aligned}
$$

The calculation for the particular case detailed above is set out below:

y	5·0	11·3	−0·4	−1·0
y		−4·3	1·4	
V	5·0	7·0	1·0	−1·0
W	—	15·6	−1·8	—
V, W, P	5·0	7·0	15·6	4·0
V, W, P	−1·0	1·0	−1·8	8·0
P, R, L	4·0	8·0	13·8	12·0
Q, S	6·0	6·0	17·4	—

$$
\begin{aligned}
6A_0 &= 12{\cdot}0,\ A_0 = 2{\cdot}0. \\
3a_1 &= 6{\cdot}0/2 + 6{\cdot}0 = 9{\cdot}0,\ a_1 = 3{\cdot}0. \\
3a_2 &= -\ 8{\cdot}0/2 + 4{\cdot}0 = 0,\ a_2 = 0. \\
3b_1 &= 0{\cdot}866 \times 13{\cdot}8 = 11{\cdot}95,\ b_1 = 4{\cdot}0. \\
3b_2 &= 0{\cdot}866 \times 17{\cdot}4 = 15{\cdot}07,\ b_2 = 5{\cdot}0.
\end{aligned}
$$

7. 24-ordinate scheme.

The process outlined above is common to all the usual methods of numerical analysis; the 24- and 48-ordinate schemes given in this and the following sections are based on the same principle.

When the variation contains no harmonics higher than the 12th (or very small components of these higher harmonics) a method of analysis utilising 24 ordinates will give the constant term and the coefficients, both sine and cosine, of the first eleven harmonics; it also gives the coefficient of the 12th harmonic cosine component.

The variation being substantially of the form

$$y = A_0 + \sum_{k=1}^{k=12} a_k \cos kx + \sum_{k=1}^{k=11} b_k \sin kx,$$

with a period of 2π radians in x, the ordinates are drawn to a graph of the function at $x = 0, 15, 30, 45 \ldots 345$ degrees, and the lengths of the ordinates are denoted, respectively, by $y_0, y_1, y_2, y_3 \ldots y_{23}$. Since the value of $\sin 12x$ is zero at all these values of x, b_{12} will not be found.

The ordinates are written down in two columns as shown at the left-hand side of Table IIA, and the sums V and differences W formed; thus $V_\iota = y_\iota - y_{24-\iota}$, and $W_\iota = y_\iota - y_{24-\iota}$. The V's and W's are then subjected to the same process, forming sums P and R and differences Q and S as shown in the table. The process is continued as shown, and the quantities, C, D, E, F, J, T, X and U determined by the relations printed at the foot of the table.

Table IIB gives the scheme for calculating the cosine coefficients a_k. For each letter reference stated in the table the corresponding value from Table IIA is substituted, *after first multiplying it by the factor in the same line under the column-heading* "(Z)." The sums of the various columns then give the values of $12a_k$ [thus, for example, $12a_2 = 0{\cdot}5000\,M_2 + 0{\cdot}8660\,M_1 + M_0$] with the exception of the 12th harmonic: $J = 24a_{12}$ instead of $12a_{12}$.

A similar procedure for completing Table IIC determines the sine coefficients in the form $12b_k$, and also the constant term:

$$F = 24\,A_0.$$

The checks printed after the table are very useful and should never be omitted. It is quite easy for slips to occur even in such simple numerical work as that required in completing the tables, and the checks enable such slips to be detected.

In applying the method, a certain amount of work can be saved when the variation is symmetrical or skew-symmetrical. If it is symmetrical, only the cosine components are present, so that

Table IIa

Analysis with 24 ordinates

y		Sum V	Diff W
0	—	0	
1	23	1	1
2	22	2	2
3	21	3	3
4	20	4	4
5	19	5	5
6	18	6	6
7	17	7	7
8	16	8	8
9	15	9	9
10	14	10	10
11	13	11	11
12	—	12	—

V		Sum P	Diff Q
0	12	0	0
1	11	1	1
2	10	2	2
3	9	3	3
4	8	4	4
5	7	5	5
6		6	—

W		R Sum	S Diff
1	11	1	1
2	10	2	2
3	9	3	3
4	8	4	4
5	7	5	5
6	—	6	—

P		Sum I	Diff M
0	6	0	0
1	5	1	1
2	4	2	2
3	—	3	—

S		K Sum	N Diff
1	5	1	1
2	4	2	2
3	—	3	—

L		Sum G	Diff H
0	3	0	0
1	2	1	1

$$C = N_1 + N_2 \qquad D = N_1 - N_2 \qquad E_0 = Q_0 - Q_4 \qquad E_1 = Q_1 - Q_3 - Q_5$$

$$F = G_0 + G_1 \qquad J = H_0 - H_1$$

$$T = M_0 - M_2 \qquad X = K_1 - K_3 \qquad U_1 = R_1 + R_3 - R_5 \qquad U_2 = R_2 - R_6$$

Table IIb

Analysis with 24 ordinates cosine coefficients

(z)	x	$2x$	$3x$	$4x$	$5x$	$6x$	$7x$	$8x$	$9x$	$10x$	$11x$	$12x$
0 2588	Q_5				Q_1		$-Q_1$				$-Q_5$	
0 5000	Q_4	M_2		H_1	Q_4		Q_4	$-G_1$		M_2	Q_4	
0 7071	Q_3		E_1		$-Q_3$		Q_3		$-E_1$		$-Q_3$	
0 8660	Q_2	M_1			$-Q_2$		$-Q_2$			$-M_1$	Q_2	
0 9659	Q_1				Q_5		$-Q_5$		.		$-Q_1$	
1	Q_0	M_0	E_0	H_0	Q_0	T	Q_0	G_0	E_0	M_0	Q_0	J
Sums of columns give $12a_k$												$24a_{12}$

TABLE IIc

Analysis with 24 ordinates—sine coefficients

(Z)	x	$2x$	$3x$	$4x$	$5x$	$6x$	$7x$	$8x$	$9x$	$10x$	$11x$	$24A_0$
0·2588	R_1				R_5		R_5				R_1	
0·5000	R_2	K_1			R_2		$-R_2$			K_1	$-R_2$	
0·7071	R_3		U_1		$-R_3$		$-R_3$		U_1		R_3	
0·8660	R_4	K_2		C	$-R_4$		R_4	D		$-K_2$	$-R_4$	
0·9659	R_5				R_1		R_1				R_5	
1	R_6	K_3	U_2		R_6	X	$-R_6$		$-U_2$	K_3	$-R_6$	F
	Sums of columns give $12b_k$											$24A_0$

Checks: (i) $y_0 = A_0 + \sum_{1}^{12} a_k$.

(ii) $\frac{1}{2}(y_1 - y_{23}) = 0·2588\,(b_1 + b_{11}) + 0·5000\,(b_2 + b_{10})$
$+ 0·7071\,(b_3 + b_9) + 0·8660\,(b_4 + b_8) + 0·9659\,(b_5 + b_7) + b_6$.

Table IIc can be dispensed with; also, all the W's are zero, so that there is no need to evaluate (as zero quantities) the R's, S's, K's, N's, C, D, X or U's. Similarly, if the variation is skew-symmetrical as it stands, the sum of all the ordinates being zero, the quantities V, P, Q, L, M, E, G, H, F, J and T are all zero, and Table IIb can be omitted. If the skew-symmetry is disguised (see p. 58) the best plan is to determine the constant term as the average of the ordinates, and then subtract it from each ordinate; the variation represented by the resulting values is then truly skew-symmetrical, with a zero constant term.

The process has been applied in full to the analysis of the variation (3.1) of Chapter V, page 124, as an illustration of the method. The numerical work is given in Table III. The final multiplications and divisions have been performed with the figures corrected to the first decimal place. The sine and cosine coefficients are converted to amplitudes and phase-angles by means of the relations

$$\left.\begin{aligned} a_k &= A_k \sin \phi_k \\ b_k &= A_k \cos \phi_k \end{aligned}\right\} \quad . \quad . \quad . \quad (7.1)$$

i.e.

$$\left.\begin{aligned} A_k^2 &= a_k^2 + b_k^2 \\ \tan \phi_k &= a_k / b_k \end{aligned}\right\} \quad . \quad . \quad . \quad (7.2)$$

In determining the value of ϕ_k from the calculated value of $\tan \phi_k$, the choice between the two possible values is dictated by a consideration of the signs of the sine and cosine coefficients; details of the method are given in Chapter I, page 15.

The numerical work is completed in Table IIID, and it is seen that the variation has the form

$$
\begin{aligned}
y = 20 &+ 5{\cdot}0 \sin (x + 105^\circ) + 3{\cdot}6 \sin (2x + 63^\circ) \\
&+ 5{\cdot}9 \sin (3x - 45^\circ) - 4{\cdot}3 \sin (5x + 120^\circ) \\
&+ 4{\cdot}0 \sin (8x - 120^\circ). \qquad (7.3)
\end{aligned}
$$

Equation (7.3) can be expressed in a slightly different form :

$$
\begin{aligned}
y = 20 &+ 5{\cdot}0 \cos (x - 15^\circ) + 3{\cdot}6 \sin 2(x + 31{\cdot}5^\circ) \\
&+ 5{\cdot}9 \sin 3(x - 15^\circ) + 4{\cdot}3 \sin 5(x + 24^\circ) \\
&+ 4{\cdot}0 \sin 8(x - 15^\circ),
\end{aligned}
$$

and if the basic variable is changed to x_1, where $x_1 = x + 15^\circ$, the expression becomes

$$
\begin{aligned}
y = 20 &- 5{\cdot}0 \cos x_1 - 3{\cdot}6 \sin 2(x_1 + 16{\cdot}5^\circ) + 5{\cdot}9 \sin 3x_1 \\
&- 4{\cdot}3 \sin 5(x_1 - 9^\circ) - 4{\cdot}0 \sin 8x_1,
\end{aligned}
$$

or

$$
\begin{aligned}
y = 20 &- 5{\cdot}0 \cos x_1 + 3{\cdot}6 \sin (2x_1 + 33^\circ) + 5{\cdot}9 \sin 3x_1 \\
&- 4{\cdot}3 \sin (5x_1 - 45^\circ) + 4{\cdot}0 \sin 8x_1.
\end{aligned}
$$

TABLE IIIA

Analysis of function (3 1) *of Chapter V, page* 124

y		Sum V	Diff W
39·4	—	39·4	—
29·4	30·0	59·4	−0·6
26·5	15·2	41·7	11·3
26·5	16·1	42·6	10·4
17·1	16·3	33·4	0·8
15·0	19·5	34·5	−4·5
12·6	27·9	40·5	−15·3
6·3	21·0	27·3	−14·7
17·7	15·6	33·3	2·1
27·6	23·7	51·3	3·9
18·3	19·5	37·8	−1·2
14·0	10·9	24·9	3·1
13·9	—	13·9	—

V		Sum P	Diff Q
39·4	13·9	53·3	25·5
59·4	24·9	84·3	34·5
41·7	37·8	79·5	3·9
42·6	51·3	93·9	−8·7
33·4	33·3	66·7	0·1
34·5	27·3	61·8	7·2
40·5	—	40·5	—

W		Sum R	Diff S
−0·6	3·1	2·5	−3·7
11·3	−1·2	10·1	12·5
10·4	3·9	14·3	6·5
0·8	2·1	2·9	−1·3
−4·5	−14·7	−19·2	10·2
−15·3	—	−15·3	—

P		Sum L	Diff. M
53·3	40·5	93·8	12·8
84·3	61·8	146·1	22·5
79·5	66·7	146·2	12·8
93·9	—	93·9	—

S		K Sum	N Diff
−3·7	10·2	6·5	−13·9
12·5	−1·3	11·2	13·7
6·5	—	6·5	—

L		Sum G	Diff H
93·8	93·9	187·7	−0·1
146·1	146·2	292·3	−0·1

C = −0·1		T = 0
D = −27·7	F = 480·0	X = 0
E_0 = 25·4	J = 0	U_1 = 36·0
E_1 = 36·0		U_2 = 25·4

TABLE IIIB

(Z)	x	$2x$	$3x$	$4x$	$5x$	$6x$	$7x$	$8x$	$9x$	$10x$	$11x$	$12x$
0·2588	1·9				8·9		−8·9				1·9	
0·5000	0·1	6·4		−0·1	0·1		0·1	−146·2		6·4	0·1	
0·7071	6·2		25·5		6·2		6·2		−25·5		6·2	
0·8660	3·4	19·5			−3·4		3·4			−19·5	3·4	
0·9659	33·3				7·0		−7·0				33·3	
1	25·5	12·8	25·4	−0·1	25·5	0	25·5	187·7	25·4	12·8	25·5	0
$12a_k$	58·0	38·7	50·9	−0·2	44·3	0	0·1	41·5	−0·1	−0·3	0	0
a_k	4·8	3·2	4·2	—	3·7		—	3·5	—	—	—	—

TABLE IIIC

(Z)	x	$2x$	$3x$	$4x$	$5x$	$6x$	$7x$	$8x$	$9x$	$10x$	$11x$	$24A_0$
0·2588	0·6				−5·0		−5·0				0·6	
0·5000	5·1	3·3			5·1		−5·1			3·3	−5·1	
0·7071	10·1		25·5		−10·1		−10·1		25·5		10·1	
0·8660	2·5	9·7		−0·1	−2·5		2·5	−24·0		−9·7	2·5	
0·9659	18·5				2·4		2·4				−18·5	
1	−15·3	6·5	25·4		−15·3	0	15·3		−25·4	6·5	15·3	480·0
$12b_k$	−15·5	19·5	50·9	−0·2	−25·4	0	0	−24·0	0·1	0·1	−0·1	A_0
b_k	−1·3	1·6	4·2	—	−2·1	—	—	−2·0	—		—	20

Checks: (i) $y_0 = 39{\cdot}4 = A_0 + \Sigma a_k$.

(ii) $\frac{1}{2}(y_1 - y_{23}) = -0{\cdot}3 = 0{\cdot}2588(b_1 + b_{11}) + 0{\cdot}5000(b_2 + b_{10})$
$+ 0{\cdot}7071(b_3 + b_9) + 0{\cdot}8660(b_4 + b_8) + 0{\cdot}9659(b_5 + b_7) + b_6$.

TABLE IIID

Amplitudes and phase-angles of harmonics found from Tables IIIB and IIIC

$[A_k \sin(kx - \phi_k) = a_k \cos kx - b_k \sin kx]$

k	a_k	b_k	A_k	$\phi_k{}^\circ$
0		–	20.0	—
1	4·8	1·3	5·0	105
2	3·2	1·6	3·6	63
3	4·2	4·2	5·9	45
4		– –	—	—
5	3·7	2·1	4·3	120
6			–	—
7	–	–	—	—
8	3.5	–2·0	4·0	120
9		– –	—	—
10	—	—	—	—
11	—	—	—	—
12	—	—	—	—

No values can be given for A_{12} and ϕ_{12} in any case, as the method gives no value for b_{12}.

The accuracy of the analysis can be judged from the fact that the variation was actually synthesised, by means of Table II in Appendix V, pages 258-260, as

$$\begin{aligned} y &= 20 + 5 \cos x_1 + 3 \sin 2x_1 + 2 \cos 2x_1 + 6 \sin 3x_1 \\ &\qquad + 3 \sin 5x_1 + 3 \cos 5x_1 + 4 \sin 8x_1 \\ &= 20 + 5 \cos x_1 + 3\cdot6 \sin(2x_1 + 34^\circ) + 6 \sin 3x_1 \\ &\qquad + 4\cdot2 \sin(5x_1 + 45^\circ) + 4 \sin 8x_1. \end{aligned}$$

8. 48-ordinate scheme.

The method of analysis utilising 48 ordinates follows precisely the same general scheme as for 24 ordinates, and is set out in Table IV. It will be seen that Table IVA is similar to Table IIA; (B) and (C) are slightly different from the corresponding parts of Table II, in that the ith and $(24 - i)$th harmonic components are grouped together, each group being derived from two columns headed "α" and "β"; the coefficients up to those of the eleventh harmonic are given by the sums of the columns α, β, and the remainder by the difference $\alpha - \beta$. These details are illustrated in Table V, which gives the numerical work for the function

$$\left.\begin{aligned} y = f(x) &= \frac{24}{\pi}x \qquad &&\text{for } -\pi/2 \leqslant x \leqslant \pi/2 \\ &= 24\left(1 - \frac{x}{\pi}\right) &&\text{for } \pi/2 \leqslant x \leqslant 3\pi/2 \\ f(x + 2\pi) &= f(x) && \end{aligned}\right\} \tag{8·1}$$

This is an example of the general form illustrated in Fig. 8c of Chapter VI, page 161, with an amplitude of 12 units instead of one unit.

TABLE IVA

Analysis with 48 ordinates

y		Sum V	Diff W
0		0	—
1	47	1	1
2	46	2	2
3	45	3	3
4	44	4	4
5	43	5	5
6	42	6	6
7	41	7	7
8	40	8	8
9	39	9	9
10	38	10	10
11	37	11	11
12	36	12	12
13	35	13	13
14	34	14	14
15	33	15	15
16	32	16	16
17	31	17	17
18	30	18	18
19	29	19	19
20	28	20	20
21	27	21	21
22	26	22	22
23	25	23	23
24	—	24	—

V		Sum P	Diff Q
0	24	0	0
1	23	1	1
2	22	2	2
3	21	3	3
4	20	4	4
5	19	5	5
6	18	6	6
7	17	7	7
8	16	8	8
9	15	9	9
10	14	10	10
11	13	11	11
12	—	12	—

W		R Sum	S Diff
1	23	1	1
2	22	2	2
3	21	3	3
4	20	4	4
5	19	5	5
6	18	6	6
7	17	7	7
8	16	8	8
9	15	9	9
10	14	10	10
11	13	11	11
12	—	12	—

P		Sum L	Diff M
0	12	0	0
1	11	1	1
2	10	2	2
3	9	3	3
4	8	4	4
5	7	5	5
6	—	6	—

S		K Sum	N Diff
1	11	1	1
2	10	2	2
3	9	3	3
4	8	4	4
5	7	5	5
6	—	6	—

L		Sum G	Diff H
0	6	0	0
1	5	1	1
2	4	2	2
3	—	3	—

N		C Sum	D Diff
1	5	1	1
2	4	2	2
3	—	3	—

$E_0 = Q_0 - Q_8$ $\quad U_1 = R_1 + R_7 - R_9$ $\quad T_0 = M_0 - M_4$

$E_1 = Q_1 - Q_7 - Q_9$ $\quad U_2 = R_2 + R_6 - R_{10}$ $\quad T_1 = M_1 - M_3 - M_5$

$E_2 = Q_2 - Q_6 - Q_{10}$ $\quad U_3 = R_3 - R_5 - R_{11}$ $\quad X_1 = K_1 - K_3 - K_5$

$E_3 = Q_3 - Q_5 - Q_{11}$ $\quad U_4 = R_4 - R_{12}$ $\quad X_2 = K_2 - K_6$

$F_0 = G_0 + G_2$

$F_1 = G_1 + G_3$

TABLE IVB

Analysis with 48 ordinates—cosine coefficients

	0		1		2		3		4		5		6		7		8		9		10		11		12	
(k)	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β
0·1305				Q_{11}								Q_7				Q_5								Q_1		
0·2588			Q_{10}			M_5					Q_2				Q_2							M_1	$-Q_{10}$			
0·3827				Q_9				E_3				$-Q_3$				$-Q_0$				E_1				Q_3		
0·5000			Q_8		M_4				H_2		Q_8				Q_8		G_2	G_1			M_4		Q_8			
0·6088				Q_7								Q_{11}				Q_1								Q_5		
0·7071			Q_6			M_3	E_2				$-Q_6$			T_1	Q_6				E_2			M_3	Q_6			
0·7534				Q_5								Q_1				$-Q_{11}$								Q_7		
0·8660			Q_4		M_2					H_1	Q_4				$-Q_4$						M_2		Q_4			
0·9239				Q_3				E_1				Q_9				Q_3				$-E_3$				Q_9		
0·9659			Q_2			M_1					Q_{10}				$-Q_{10}$							M_5	Q_2			
0·9914				Q_1								$-Q_5$				Q_7								$-Q_{11}$		
1	F_0	F_1	Q_0		M_0		E_0		H_0		Q_0		T_0		Q_0		G_0	$-G_1$	E_0		M_0		Q_0		H_0	H_2
Σ_α Σ_β																										
Sum Diff.	$48A_0$ $48a_{24}$		$24a_1$ $24a_{23}$		$24a_2$ $24a_{22}$		$24a_3$ $24a_{21}$		$24a_4$ $24a_{20}$		$24a_5$ $24a_{19}$		$24a_6$ $24a_{18}$		$24a_7$ $24a_{17}$		$24a_8$ $24a_{16}$		$24a_9$ $24a_{15}$		$24a_{10}$ $24a_{14}$		$24a_{11}$ $24a_{13}$		— $24a_{12}$	

TABLE IVc

Analysis with 18 ordinates—sine coefficients

	1		2		3		4		5		6		7		8		9		10		11		12	
(S)	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β	α	β
0·1305	R_1								$-R_5$				R_7								$-R_{11}$			
0·2588		R_2	K_1							R_{10}				R_{10}					K_5			R_2		
0·3827	R_3				U_1				R_9				R_3				U_3				R_9			
0·5000		R_4		K_2			C_1			R_4				R_4						K_4		$-R_4$		
0·6088	R_5								R_1				R_{11}								$-R_7$			
0·7071		R_6	K_3			U_2				R_6	X_1			R_6				U_2	$-K_3$			R_6		
0·7934	R_7								R_{11}				R_1								R_5			
0·8660		R_8		K_4				C_2		$-R_8$				R_8	D_1	D_2				$-K_2$		$-R_8$		
0·9239	R_9				U_3				R_3				R_9				U_1				$-R_3$			
0·9659		R_{10}	K_5							R_2				R_2					K_1			R_{10}		
0·9914	R_{11}								R_7				R_5								R_1			
1		R_{12}		K_6		U_1	C_3			R_{12}		X_2		R_{12}				$-U_1$		K_6		$-R_{12}$	C_1	C_3
Σ_α																								
Σ_β																								
Sum	$24b_1$		$24b_2$		$24b_3$		$24b_4$		$24b_5$		$24b_6$		$24b_7$		$24b_8$		$24b_9$		$24b_{10}$		$24b_{11}$		—	
Diff.	$24b_{23}$		$24b_{22}$		$24b_{21}$		$24b_{20}$		$24b_{19}$		$24b_{18}$		$24b_{17}$		$24b_{16}$		$24b_{15}$		$24b_{14}$		$24b_{13}$		$24b_{12}$	

Checks: (i) $y_0 = A_0 + \Sigma a_k$.

(ii) $\frac{1}{2}(y_1 - y_{17}) = 0{\cdot}1305(b_1 + b_{23}) + 0{\cdot}2588(b_2 + b_{22}) + 0{\cdot}3827(b_3 + b_{21}) + \ldots\ldots - 0{\cdot}9914(b_{11} - b_{13}) + b_{12}$.

TABLE VA

Analysis of function (8 1)

u		Sum V	Diff W	v		Sum P	Diff Q	p		Sum I	Diff M
0											
1	1		2								
2	− 2		4								
3	− 3		6					0	0	0	0
4	4		8								
5	5		10								
6	6		12	0	0	0	0				
7	7		14								
8	− 8		16								
9	− 9		18								
10	− 10		20								
11	− 11		22					0	0	0	0
12	−12	0	24								
11	11		22								
10	10		20								
9	− 9		18	2	2	4					
8	− 8		16	4	4	8		S		K Sum	N Diff
7	− 7		14	6	6	12					
6	− 6		12	8	8	16					
5	− 5		10	10	10	20		L		Sum G	Diff H
4	− 4		8	12	12	24	0				
3	− 3		6	14	14	28					
2	− 2		4	16	16	32					
1	− 1		2	18	18	36		0	0	0	0
0	—		0	20	20	40					
				22	22	44					
				24	—	48					
				W		R Sum	S Diff	0	0	0	0
								N		C Sum	D Diff

$$\left.\begin{matrix} E_0 \\ E_1 \\ E_2 \\ E_3 \end{matrix}\right\} = 0 \qquad \begin{aligned} U_1 &= -4 \\ U_2 &= -8 \\ U_3 &= -12 \\ U_4 &= -8 \end{aligned} \qquad \left.\begin{matrix} T_0 \\ T_1 \\ X_1 \\ X_2 \end{matrix}\right\} = 0$$

$$\left.\begin{matrix} F_0 \\ F_1 \end{matrix}\right\} = 0$$

TABLE VB

Sine coefficients

	1		3		5		7		9		11	
(l)	α	β	α	β	α	β	α	β	α	β	α	β
0 1305	0 52				—2 61		3 65				— 5 74	
0 2588		2 07				10 35		10 35				2 07
0 3827	4 59		— 1 53		—13 78		4 59		4 59		13 78	
0 5000		8 00				8 00		— 8 00				— 8 00
0 6088	12 18				2 44		—26 79				— 17 05	
0 7071		16 97		—5·66		— 16 97		16 97		—5 66		16 97
0·7934	22 22				34 91		3 17				15 87	
0 8660		27 71				—27 71		27 71				— 27 71
0 9239	33 26		— 11 09		11 09		33 26		— 3 70		—11 09	
0 9659		38 64				7 73		7·73				38 64
0 9914	43 62				—27 76		— 19 83				3 97	
1		24 00		—8 00		24 00		—24 00		8 00		—24 00
Σ_α	116·39		12·62		4·29		—1·95		0·89		—0·26	
Σ_β	117·39		—13·66		5 40		—3 18		2 34		—2·03	
Sum	233 78		26 28		9 69		—5 13		3 23		—2 29	
Diff.	1·00		1·04		1 11		1 23		—1 45		1·77	
	9 74		—1·09		0·40		—0·21		0·13		—0·10	
b_k	0 04		0·04		—0 05		0 05		—0 06		0 07	

The wave being skew-symmetrical, all the cosine terms are zero; and since the wave is also alternant, the even sine terms are zero. Much of the work is therefore simplified. The odd sine terms are determined in Table VB, and are reprinted in the second column of Table VI. These figures are corrected to the second decimal

TABLE VI

Comparison of approximate and correct values of coefficients for the function (8.1)

k	b_k Numerical	b_k Accurate
1	9·74	9·73
3	−1·09	−1·08
5	0·40	0·39
7	−0·21	−0·20
9	0·13	0·12
11	−0·10	−0·08
13	0·07	0·06
15	−0·06	−0·04
17	0·05	0·03
19	−0·05	−0·03
21	0·04	0·02
23	−0·04	−0·02

place. In the last column of the table the accurate values are given, also correct to two decimal places. These values were determined by means of the formulæ (5.5) of Chapter VI, page 162, multiplied by 12 on account of the modified amplitude. The effect of neglecting the harmonics higher than the 23rd can be seen from Table VI, and it will be observed that the values calculated by the numerical method are *not* correct to the second decimal place.

9. Practical notes on the method.

When separate observations are made to enable values of a periodic variation to be measured for various values of the basic variable, it is desirable to arrange matters so that the measurements can themselves be used as the ordinates for a scheme of analysis. For this purpose the values of the basic variable should be distributed evenly over the cycle, and in order to make use of the available computation schemes the number of the ordinates should be 12, 24, 48, or 72.

When it is not possible to fulfil these conditions, and the values

of the variable have to be determined for some unusual number of values of the basic variable, or for some such number of values distributed unevenly over the cycle, it is preferable *not* to interpolate numerically or graphically (i.e. by making use of finite differences or by drawing a smooth curve through points representing the known values) but to proceed as follows. Suppose that there are N distinct values of the function y, denoted by y_r, corresponding to the N values x_r of the basic variable x ($r = 1, 2 \ldots$ N). If N is odd, the function may be approximated by the series

$$\left.\begin{aligned} &A_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \\ \text{where} \qquad &2n + 1 = N \end{aligned}\right\} \quad . \quad . \quad (9.1)$$

and if N is even, by the series

$$\left.\begin{aligned} &A_0 + \sum_{k=1}^{n-1} (a_k \cos kx + b_k \sin kx) + a_n \cos nx \\ \text{where} \qquad &2n = N \end{aligned}\right\} \quad . \quad (9.2)$$

Substituting in turn $x = x_1, x_2 \ldots x_N$ in the series (9.1) or (9.2), and equating the sum of the series to $y_1, y_2 \ldots y_N$, there are obtained N equations in N unknown quantities, and these unknown quantities (the Fourier coefficients) can be found by any of the usual methods of solving simultaneous equations.

In the case of recorded waveforms it is usually necessary to enlarge the graph by optical projection or photographic means, so as to enable the ordinates to be measured with a sufficient degree of accuracy. The division of the cycle into 12, 24, etc., equal spaces may conveniently be performed on a drawing-board by the familiar method of elementary geometry; this will be found to give more accurate results than by measurement along the axis of multiples of $\lambda/24$, say, where λ is the wavelength. On the other hand, the record may be enlarged so that the wavelength becomes a convenient length, such as 9·6 inches, which can easily be divided into 12, 24 or 48 equal spaces by means of an inch scale graduated in tenths.

When it is desired merely to determine the coefficients of a single harmonic, or of two or three harmonics, the full scheme of computation need not be completed. If, for example, it is required to evaluate the seventh harmonic in a 48-ordinate analysis, it is sufficient to find the quantities V, W, Q and R in Table IVA, and to use the columns (7) in Tables IVB, C.

10. Proof of summation formulæ.

The formulæ (5.4) on page 193, which are used in the solution of the equations (5.3), are here proved.

Lemma.
$$\sum_{r=1}^{N} \cos rz = \frac{\cos (N+1)\frac{z}{2} \sin N\frac{z}{2}}{\sin \frac{z}{2}}, \qquad (10.1)$$

where z is any quantity.

For, the summations extending from $r = 1$ to $r = N$,

$$\sin \frac{z}{2} \Sigma \cos rz = \Sigma \cos rz \sin \frac{z}{2},$$
$$= \tfrac{1}{2} \Sigma[\sin (r + \tfrac{1}{2})z - \sin (r - \tfrac{1}{2})z],$$
$$= \tfrac{1}{2}\left[\sin (N + \tfrac{1}{2})z - \sin \frac{z}{2}\right],$$

whence the result.

Proof. The summations covering the same extent as before,

$$\Sigma \sin mr\frac{2\pi}{N} \sin nr\frac{2\pi}{N} = \tfrac{1}{2}\Sigma\left[\cos (m-n)r\frac{2\pi}{N} - \cos (m+n)r\frac{2\pi}{N}\right],$$

$$= \frac{1}{2}\left[\frac{\cos (m-n)(N+1)\frac{\pi}{N} \sin (m-n)\pi}{\sin (m-n)\pi/N} - \frac{\cos (m+n)(N+1)\frac{\pi}{N} \sin (m+n)\pi}{\sin (m+n)\pi/N}\right]. \qquad (10.2)$$

Now if $m \neq n$, $m + n < N$, $\Sigma = 0$.

If $m = n = k$, $\cos (N + 1)(m - n)\pi/N = 1$,

and
$$\underset{m \to n}{\mathrm{Lt.}} \frac{\sin(m-n)\pi}{\sin(m-n)\pi/N} = N.$$

And if $2k \neq N$, the second term in the square brackets is zero; hence in this case $\Sigma = N/2$.

If $m = n = k = N/2$, $\sin^2 kr\frac{2\pi}{N} = \sin^2 r\pi = 0$, so that if $m = n$ the results are obtained :

$$\left.\begin{aligned}\sum_{r=1}^{N} \sin^2 kr\frac{2\pi}{N} &= N/2 \quad \text{if} \quad k \neq N/2\\ &= \quad 0 \quad \text{if} \quad k = N/2\end{aligned}\right\} \qquad (10.3)$$

In a similar manner it may be proved that

$$\left.\begin{aligned}\sum_{r=1}^{N} \cos mr\frac{2\pi}{N}\begin{pmatrix}\sin\\ \cos\end{pmatrix} nr\frac{2\pi}{N} &= 0 \quad \text{if} \quad m \neq n, m+n < N \\ &= \begin{pmatrix}0\\ N/2\end{pmatrix} \quad \text{if} \quad m = n \neq N/2 \\ &= \begin{pmatrix}0\\ N\end{pmatrix} \quad \text{if} \quad m = n = N/2\end{aligned}\right\}. \tag{10.4}$$

The equations (10.2, 3, 4) are the same as (5.4), which were to be established.

CHAPTER VIII

MECHANICAL AND OTHER AIDS TO ANALYSIS

1. Introductory.

In certain instances the process of determining the Fourier coefficients for a function or recorded waveform can be performed by a mechanical, optical or electrical instrument designed specially for the purpose. New varieties of such instruments are continually being devised, both for general and specific applications, and it is impracticable to attempt a comprehensive description of all the available arrangements. Neither is so complete a description desirable, several important varieties of instrument being very limited in their range of useful employment.

One disadvantage of mechanical and other analysing instruments lies in the fact that the time-base of the record must be accurately constant; if the record does not fulfil this requirement it may be possible to replot it on a constant time-base, but the process is tedious.

In this chapter one general-service type of mechanical analyser is described in detail—the Harvey machine. Several other types of instrument receive brief mention, and some notes are included on the use of filter circuits. The chapter concludes with a reference to the scientific computing services, which are prepared to undertake numerical analyses.

2. Mechanical analysers

There are several different types of mechanical analyser for general use. Many of these are described in the Handbook of the Napier Tercentenary Exhibition (reference 1 in the Bibliography). A typical machine is the Harvey Harmonic Analyser, which is available in several forms: the most complete instrument will give the Fourier coefficients of the first 14 harmonics, and can also be used as a straightforward planimeter or integrator. The graph of the function to be analysed must be drawn so that its wavelength is 24 cm.; recorded waveforms can of course be adjusted photographically to fulfil this condition. The following description by the inventor, Mr. J. Harvey, A.R.C.S., B.Sc., is reprinted from "Engineering," December 21, 1934, by permission of the publishers.

Mr. Harvey's notation is in general the same as that adopted in this book, the only important exceptions being the use of the

symbols "a_0" and "n" to represent the constant term and the harmonic reference number (A_0 and k in this book).

Harvey Harmonic Analyser

(By J. Harvey, A.R.C.S., B.Sc.; reprinted from "Engineering," December 21, 1934, by permission of the publishers.)

In the *Philosophical Magazine* of April, 1895, Mr. G. Udny Yule described a principle and an instrument based on the principle by which, with the employment of a planimeter, the harmonic analysis of a function could be effected. The author of this article had rediscovered the principle, and had almost constructed the instrument about to be described before he became aware of Mr. Yule's paper. This instrument is independent of a planimeter; further, the mechanism is adapted for finding area, and first and second moments of area about an axis. The explanation of the principle will be given in a modified form from that in which Mr. Yule gave it, so as to bear more directly on the special instrument.

The formulæ for the coefficients, a_0, a_n, b_n, $n = 1, 2, \ldots$ of a Fourier series, $a_0 + \Sigma a_n \cos nx + \Sigma b_n \sin nx$, for a function $f(x)$ of an angle x, from $x = 0$ to $x = 2\pi$, are

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx$$

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx,$$

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\,dx.$$

The coefficient a_0 is the mean height of the curve $y = f(x)$ over the range $0 - 2\pi$. It may be found by dividing the area bounded by the curve, the x-axis and the ordinates $x = 0$ and $x = 2\pi$, by the length a of the base line representing the range $0 - 2\pi$. The method of employing the instrument for finding area will be described later.

The theory on which the instrument finds a_n and b_n will now be explained.

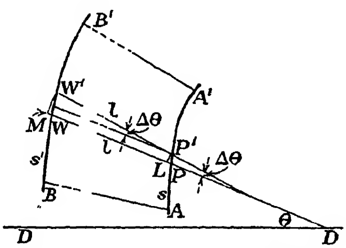


FIG. 1.—Theory of analyser ("Engineering").

Let the extremities P and W of a bar of length l, move over curves AA′ and BB′ from an initial position AB of the bar to a final position A′B′, Fig. 1.

Let PW be a position of the bar which makes an angle θ with a fixed line DD, and such that the arc AP $= s$, and the arc BW $= s'$. Let P′W′ be a position of the bar near to PW and making with it an angle $\Delta\theta$.

If P′L and W′M are perpendiculars on PW,

$$\text{LP}' = \text{chord PP}' \,.\, \sin \text{LPP}',$$
$$\text{MW}' = \text{chord WW}' \,.\, \sin \text{MWW}',$$

also,

$$\text{MW}' = \text{LP}' + \text{W}'\text{P}' \sin \Delta\theta :$$

from which,

$$\text{chord WW}' \,.\, \sin \text{MWW}' = \text{chord PP}' \,.\, \sin \text{LPP}' + \text{W}'\text{P}' \,.\, \sin \Delta\theta,$$

or,

$$\frac{\text{chord WW}'}{\Delta s'} \sin \text{MWW}' \,.\, \Delta s' = \frac{\text{chord PP}'}{\Delta s} \,.\, \sin \text{LPP}' \,.\, \Delta s + \text{W}'\text{P}' \,.\, \sin \frac{\Delta\theta}{\Delta\theta} \,.\, \Delta\theta$$

Integrating between the positions AB and A′B′, we have

$$\int_{\text{B}}^{\text{B}'} \sin \phi' ds' = \int_{\text{A}}^{\text{A}'} \sin \phi ds + l(\Theta' - \Theta),$$

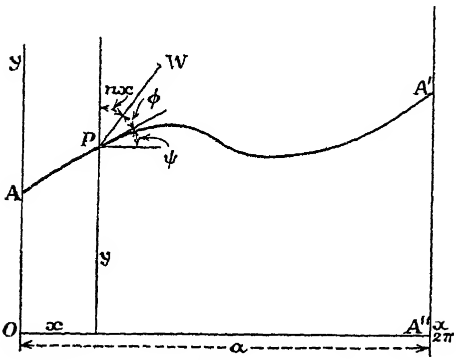


FIG. 2.—Theory of analyser (“ Engineering ”)

where ϕ and ϕ' are the angles which the tangents to the curve AA′ and BB′ at P and W, respectively, make with PW, and Θ and Θ' are the angles which AB and A′B′ respectively make with DD.

If, during the motion, there is a wheel fixed to the bar at W, which can turn freely about its axis coincident with PW, and which can roll and slide on the paper, the distance w, rolled by the wheel is

$$\int_{\text{B}}^{\text{B}'} \sin \phi' \,.\, ds'.$$

Hence
$$w = \int_{\text{A}}^{\text{A}'} \sin \phi ds + l(\Theta' - \Theta).$$

Let, Fig. 2, the function $y = f(x)$ between $x = 0$ and $x = 2\pi$ be represented by the curve AA′, the range of angle being represented by the length a. Let one end P of a bar PW, with wheel at W as described above, be brought round the boundary OAA′A″O. Let PW during the movement make, for any abscissa x of P, an angle nx with the positive direction of y, n being a positive integer; so that PW makes n complete revolutions clockwise as P traces AA′, and makes n complete revolutions anticlockwise as P traces A″O. As P moves along the lines OA and A′A″, PW lies along these lines, and the wheel slides only on the paper. As P moves along the curve AA′, PW makes with the tangent to the curve at P an angle $\phi = \frac{\pi}{2} - (nx + \psi)$, where ψ is the angle which the tangent to the curve makes with the x-axis. The distance, w_1, rolled by the wheel positively, in the movement is

$$w_1 = \int_{x=0}^{x=2\pi} \sin \phi ds + l \,.\, 2\pi n,$$

$$= \int_{x=0}^{x=2\pi} \cos (nx + \psi) ds + l \,.\, 2\pi n,$$

$$= \int_{x=0}^{x=2\pi} \cos nx \cos \psi ds - \int_{x=0}^{x=2\pi} \sin nx \sin \psi ds + l \,.\, 2\pi n$$

$$= \frac{a}{2\pi}\int_{0}^{2\pi} \cos nx \, dx - \int_{x=0}^{x=2\pi} \sin nx \, dy + l \,.\, 2\pi n$$

$$\left[\text{Since } \cos \psi = \frac{a}{2\pi}\frac{dx}{ds}, \text{ and } \sin \psi = \frac{dy}{ds}\right]$$

$$= -\left[y \sin nx - n\int y \cos nx \, dx\right]_{x=0}^{x=2\pi} + l \,.\, 2\pi n.$$

[By the rule for Integration by Parts.]

$$= n\int_{0}^{2\pi} y \cos nx \, dx + l \,.\, 2\pi n.$$

As P moves along the path A″O, the distance, w_2, rolled negatively by the wheel is, since y is zero for a point on A″O, equal to $l \,.\, 2\pi n$.

Hence in the complete circuit by P, the distance rolled positively is

$$w_1 - w_2 = n\int_{0}^{2\pi} y \cos nx \, dx.$$

If d is the diameter of the rolling wheel and m is the number of turns made by the wheel, as given by a counter attached,

$$w_1 - w_2 = m\pi d.$$

Hence $$m\pi d = n\int_{x=0}^{x=2\pi} y \cos nx\, dx,$$

from which $$\frac{md}{n} = \frac{1}{\pi}\int_0^{2\pi} y \cos nx\, dx = a_n,$$

a cosine coefficient in the Fourier series.

Let, during the movement of P round the path OAA′A″O, the bar PW make, for an abscissa x of P, the angle nx with the negative direction of x. As P moves over the line OA, the wheel rolls positively a distance $w_1 =$ OA.

As P moves over the line A′A″, the wheel rolls negatively a distance $w_2 =$ A′A″.

As P moves over the curve AA′, PW makes with the tangent to the curve at P an angle $\phi = \pi - (nx + \psi)$; the distance w_3 rolled positively by the wheel is

$$\begin{aligned} w_3 &= \int_{x=0}^{x=2\pi} \sin \phi ds + l \,.\, 2\pi n, \\ &= \int_{x=0}^{x=2\pi} \sin (nx + \psi) ds + l \,.\, 2\pi n \\ &= \int_{x=0}^{x=2\pi} \sin nx \cos \psi ds + \int_{x=0}^{x=2\pi} \cos nx \sin \psi ds + l \,.\, 2\pi n, \\ &= \frac{a'}{2\pi}\int_{x=0}^{x=2\pi} \sin nx\, dx + \int_{x=0}^{x=2\pi} \cos nx\, dy + l \,.\, 2\pi n, \\ &= \Big[y \cos nx + n\int y \sin nx\, dx \Big]_{x=0}^{x=2\pi} + l \,.\, 2\pi n, \\ &= \text{A}''\text{A}' - \text{OA} + n\int_{x=0}^{x=2\pi} y \sin nx\, dx + l \,.\, 2\pi n. \end{aligned}$$

As P moves over the path A″O, the distance w_4 rolled negatively by the wheel is $l \,.\, 2\pi n$.

Hence in the complete circuit by P of the boundary, the wheel rolls positively a distance $w_1 + w_3 - w_2 - w_4$

$$= n\int_{x=0}^{x=2\pi} y \sin nx\, dx.$$

If m is the number of revolutions of the wheel, then

$$\frac{md}{n} = \frac{1}{\pi}\int_0^{2\pi} y \sin nx\, dx = b_n,$$

a sine coefficient in the Fourier series.

The instrument is designed so that one end P of a bar PW can be made to trace any boundary within a range a representing 2π

radians in the x-direction, and so that PW can make the angle nx with the positive y-direction or the negative x-direction for co-ordinates x and y of P, according as a_n or b_n is required, from $n = 1$ to $n = 6$.

A rectangular frame CDEF, Fig. 3, can be moved in the y-direction on three rollers, R_1, R_2, and R_3, two of which, R_1 and R_2, have the same axle, to which and also to the x-axis the axis of R_3 is parallel. A bar BB is fixed to the frame. On its top surface is cut a groove

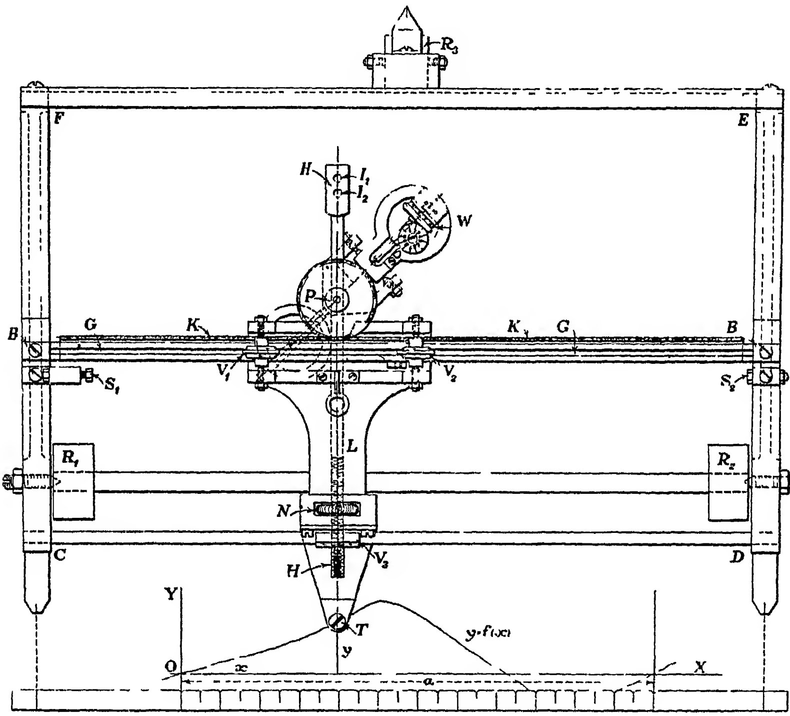


FIG. 3.—Diagram of analyser (" Engineering ").

GG, and to one vertical face is fixed a rack KK; both groove and rack are parallel to the x-direction. A trolley L is suspended by two edged wheels V_1 and V_2, which roll on the edges of the groove, and a wheel V_3 which rolls on the bar CD. The trolley can be moved freely in the x-direction between stops S_1 and S_2, so that a tracing point T fixed to it can traverse a distance a (= 24 cm.) in the x-direction, representing the range $0 - 2\pi$. Thus, owing to the freedom of motion of the frame in the y-direction and of the trolley in the x-direction, the point T can be made to trace any

boundary within the range $0 - 2\pi$, in the x-direction; there is no limit to the movement of T in the y-direction.

A bar HH in the y-direction can be drawn through guides under the trolley by a milled screw-nut N. In HH there is a vertical hole in which a spindle P can turn. To the upper end of the spindle one of a set of gear wheels of radii $\frac{a}{2\pi n}$, $n = 1, 2, \ldots 6$, can be fixed. To the lower part of the spindle P is fixed a small frame in which is set the recording wheel W to roll on the paper, with its axis passing through the axis of the spindle P.

To set the instrument, the trolley is brought up to the stop S_1, so that T is at the origin of co-ordinates, then by means of the screw-nut N, the gear wheel is drawn in to engage with the rack and so that the direction of PW is the positive direction of y or the negative direction of x, according as a_n or b_n is required.

As T is brought round the boundary, P will describe a similar boundary, and PW will make the angle nx with the positive y- or the negative x-direction for a position (x, y) of T. The diameter d of the rolling wheel W is 2 cm. The reading m of the wheel when multiplied by $\frac{n}{2}$ gives the required coefficients as centimetres for the plotted curve.

The instrument is manufactured by Messrs. A. J. Amsler and Company, Schaffhausen, Switzerland.

3. Other forms of analysing instruments.

Many instruments have been devised to perform the necessary integrations by other than purely mechanical methods. Two of these are briefly described below.

Cathode-ray tube method. The periodic variation which is required to be analysed is converted, by some suitable method, into a voltage wave which is applied to one pair of deflector plates in a cathode-ray tube. A vertical excursion of the light-spot on the screen is thus produced, this excursion being proportional to the variation being analysed. To a second pair of deflector plates is applied a constant amplitude sinusoidal voltage variation, which imparts to the light-spot a sinusoidal horizontal motion. When the frequency of the horizontal or time-base voltage is suitably related to the fundamental frequency of the variation, a Lissajou figure is produced, the area of which is proportional to the Fourier coefficient of a particular harmonic component. In practice, arrangements are made so that the time-base frequency is easily adjusted to any exact multiple of the fundamental frequency of the complex variation,

and so that the correct phase-relationship is obtained. A series of Lissajou figures is obtained, from which the Fourier coefficients are determined by means of a polar planimeter.

A more detailed description of the instrument is given in reference 2 in the Bibliography at the end of the book, where it is stated that further development of the method may enable the hitherto difficult analysis of ultra-high-frequency waves to be performed with comparative ease.

Optical integration method. An ingenious method whereby the analysis is performed optically is described in reference 3 in the Bibliography. A variable-area film record is made of the variation to be analysed. This is similar to a normal waveform, except that a line of demarcation between opaque and transparent parts of the film replaces the usual opaque line on a transparent background; the effect can be obtained from an ordinary record by "blacking-in" the area to one side of the trace. [Engineers familiar with the R.A.E. Vibrograph will notice that the record produced by the vibrograph is of this type.]

The instrument is provided with 30 "harmonic plates," each of which comprises a variable-density sinusoidal record. The plate corresponding to a particular harmonic is set up parallel with the variable-area record, and as the distance between the two records is varied a photo-cell automatically indicates the values of the corresponding harmonic. The difference between the maximum and minimum readings determines the amplitude, and the position of the harmonic plate corresponding to a maximum photo-cell reading gives the phase-angle. The instrument is so arranged that curves showing the first 30 harmonics of a complex variation can be obtained in 90 seconds.

The instrument was devised in the Bell Telephone Research Laboratories of America, for a specific purpose, and so far as is known at the time of writing it is not available as a commercial product.

4. Filter circuits.

The electrical type of waveform analyser, utilising a highly selective filter circuit, is capable of giving accurate and useful results in certain restricted ranges of application. The variation to be analysed is, so far as the instrument is concerned, a voltage variation which in some forms of design is "mixed" with the output of a local oscillator to set up sum and difference frequencies. The following particulars of an instrument of this type are quoted from a Catalogue of the General Radio Company of America (reference 4 in the Bibliography):

Type : 736A, Wave Analyser.
Frequency range : 20 to 1600 C.P.S.
Selectivity : approximately 4 C.P.S. "flat top" to the response curve. Response at 5 C.P.S. from peak is about 1/30 (15 decibels down), and at 10 C.P.S. from peak about 1/1000 (30 decibels down).
Frequency calibration : $\pm$ 2 per cent.
Voltage accuracy : within $\pm$ 5 per cent.
Dimensions : $19\frac{1}{2} \times 25\frac{1}{8} \times 10\frac{7}{8}$ in. overall.
Weight : 85 lb.

Thus with this particular instrument set at 100 C.P.S., the voltmeter will give the average value of the voltage of all frequencies within the band 98-102 C.P.S.; but the frequency calibration figure must not be neglected. The actual range selected may be from 96-100 to 100-104 C.P.S., so that at this setting there is an uncertainty to the extent of $\pm$ 4 per cent. At 1000 C.P.S. the instrument selects the band of frequencies from 998 to 1002 C.P.S., and the possible error in the frequency calibration may shift this band to 978-982 or to 1018-1022 C.P.S. With a well-defined harmonic series this measure of uncertainty is unimportant; but in certain applications, particularly in the analysis of vibrations arising simultaneously from two sources with different fundamental frequencies, the error may render the method inapplicable. This point is considered below.

The practical details of the application of a simpler form of filter arrangement to the analysis of torsional vibrations in internal combustion engines has been given by Stansfield (reference 5 in the Bibliography).

There can be no doubt that in such an application the use of filters is very convenient and yields valuable results, the frequencies of the various harmonics being fairly widely separated.

One field of study in which the use of filters has periodically been suggested is the analysis of records showing the fluctuating stresses in aircraft propeller blades under operating conditions. In this study many hundreds of records are taken in each test, and an apparatus capable of giving accurate results rapidly would be an inestimable boon. Two methods of application have been suggested : (i) the abolition of the records, the filters being used to analyse the output from the strain gauges, and (ii) the conversion of the recorded waveforms into electrical variations which would then be analysed by the filters. In the latter method great difficulty would be experienced in the initial conversion of the recorded

waveforms into electrical variations without distortion of the time-base representation. Furthermore, there are certain disadvantages inherent in the general nature of the filter device which prevent its use from being much help in this connection :

(i) Suppose the engine to have a reduction gear-ratio 0·252 : 1. A vibration which occurs four times per propeller revolution, i.e. whose frequency is four times propeller R.P.M. or 1·008 times engine R.P.M., is easily confused with the engine-excited vibration whose frequency is equal to the engine R.P.M. ; the difference in the two frequencies is less than 1 per cent., and in order to distinguish between the two a very high degree of accuracy in the frequency calibration and discrimination would be required. Given a recorded waveform of sufficient length, frequency discrimination of this order is quite easy.

(ii) With the simpler type of filter the calibration constant (relating the indicated reading to stress value) varies with the frequency. With such an instrument it would not be possible to obtain a direct reading of stress values on the dial, and use would have to be made of a conversion table or graph.

(iii) A predominant component, sufficiently widely separated in frequency from other components to enable filters to be employed satisfactorily, can easily be picked-out by the envelope method of inspection analysis.

(iv) A filter circuit employing condensers in the operative part would require frequent calibration, as condensers are notoriously unreliable. An accurate frequency source is required for the calibration, and this in itself is not easy to obtain. Very high accuracy in frequency determination is required in order to distinguish engine and propeller orders in certain cases, as instanced in (i) above, and it would always be desirable to obtain a permanent recorded waveform to ensure that no error was made in determining the source of the vibration.

(v) Even if it is possible, by refinement of design, to obtain the requisite degree of accuracy in frequency calibration, the unavoidable fluctuation in running speed of the engine would completely upset the analysis. In this connection it may be noted that S. Timoshenko raises the same objection to the use of a comparatively accurate mechanical frequency-meter, which is not even as selective as modern electric filters (reference 6 in the Bibliography). On the other hand, with a recorded waveform which incorporates adequate time and revolution markings, this speed fluctuation causes no difficulty.

(vi) In strain-gauge and other tests the phase-relationship between simultaneous vibrations in various parts of the propeller are of importance, and the normal filter arrangement affords no information on this point.

(vii) In many tests the important measurement is that of the overall excursion in a complex wave whose wavelength may be quite long. Even if the amplitude of every harmonic present is found and measured separately by means of the filters, the overall amplitude of the waveform cannot be determined by mere addition ; in cases where there are two components of frequency ratio 2 : 1 or 3 : 1, for example, the total variation is often less than the sum of the individual double amplitudes, the discrepancy depending upon the phase-relationship.

The above remarks should not be taken as condemnatory of the general notion of employing filters for waveform analysis, but refer only to a specific application. There is no doubt that in other fields of study, particularly in acoustic research, filters provide the only satisfactory method of attack.

5. Computing service.

When it is required to perform a large number of harmonic analyses of the type which is easily handled by the numerical method described in Chapter VII, the burden of carrying through the tedious calculation may be transferred to an organisation equipped with machines specially designed for such purposes. These concerns have at their disposal a large number of calculating machines of various types, and a tremendous background of experience in all forms of computation. In particular, harmonic analysis by numerical methods can be performed on the Hollerith punched card machine. The practice of deputing the routine work of numerical analysis, or indeed of any variety of calculation, to a highly-trained and well-equipped staff of expert computers has much to recommend it, and is being adopted to an increasing extent. The engineer or scientist is rarely able to perform routine calculations in the speediest and most efficient manner, as he lacks the training and in any case has other matters to attend to ; and although many research departments include amongst their numbers highly trained mathematicians, it must be remembered that mathematical knowledge is not synonymous with computing skill. Where numerical harmonic analysis is concerned, the use of the Hollerith machine expedites the computation to an amazing extent, and as these machines cannot be hired for short periods of time there is every advantage in passing on the work to a firm which has the use of one. Such a firm exists in Great Britain (see reference 7 in the Bibliography).

CHAPTER IX

PRACTICAL REQUIREMENTS FOR WAVEFORMS

1. Introductory.

In the present chapter are discussed the more important practical details of waveform recording, attention to which will render the task of analysis much easier.

The general technique of obtaining records, with the attendant problems of instrumentation, lies outside the scope of this book. Practical hints on the use of recording torsiographs are given by Ker Wilson, reference 1 in the Bibliography. Cathode-ray tube technology is treated in the standard textbooks on that subject, and information concerning the use of various types of vibrographs is given in reference 2.

The matters here discussed pertain to the analysis of recorded waveforms: clarity and general cleanliness of the records with reference to the various recording media, linearity of the overall response of the recording system, and the provision of adequate length of record and time-reference markings are considered in turn.

2. Clarity of records—various media.

A good analyst can obtain almost as much and as accurate information from a dirty, scrappy record as from a clean and clear trace; but it is not fair to expect him to maintain this standard in the routine analysis of hundreds of records. Faint or blurred traces lead to errors, apart from the imposition of considerable eye-strain.

The traces produced by marking-pens are usually clear and bold. Care must, however, be taken to avoid blurring, which may result from two causes. First, the ink may run between the fibres of the paper on which the record is made, causing a general "fuzziness." This is illustrated at (*a*) in Fig. 1. The main reasons for the occurrence of this fault are: incorrect ink or paper, or maladjustment of the ink feed to the pen. The maker's instructions should be studied carefully to correct the fault. One serious result of the fuzziness is the difficulty of measuring amplitudes accurately, the thickness of the trace being indeterminate. Secondly, the ink may "blot-over" or smudge, as illustrated at (*b*) in the diagram. This fault may be caused by using the wrong ink (special quick-drying

varieties being best), having too free a flow of ink to the pen, or not having adequate drying arrangements. Some instruments using ink incorporate a blotting roller, beneath which the trace passes before there is any possibility of smudging on some metal part of the instrument.

Pencil traces are usually most unsatisfactory. In order to reduce the bearing pressure, and hence the drag on the recorder, to a minimum a fairly soft pencil must be used, and this entails a diminution in permanency of the trace and a rapid wearing-away of the pencil-point. Pencil records are much more easily blurred and made "grubby" by careless handling than are properly dried ink records, and in fact there appears to be no redeeming feature to recommend the use of pencil recording points.

A silver or brass stylus may be used to make traces on leaded paper. Such a stylus is superior to a pencil-point, since it is much harder, but the records are liable to fade after some time and so are not very convenient if records must be kept for future reference.

Certain recording instruments produce traces on celluloid strips, the surface of the celluloid being dented by plastic deformation, not by a scratch. Such records are not suitable for direct analysis, but should first be photographed. It may be remarked that the shape of the indentation has been carefully designed by the Cambridge Instrument Company, so that the photographic copy made by transmitted parallel light is very clear.

Very neat records are obtainable on waxed paper. The original paper of this type was manufactured in Germany, and consisted of a bright red backing coated with white wax. The recording stylus scratched the wax off the paper and left a very clear red trace, where the paper-backing showed through. A very good substitute is available in this country, although the contrast is not quite as great as was the case with the German paper. One disadvantage of all waxed paper records is that subsequent mishandling leaves indelible marks ; any scratch or undue pressure on the surface removes wax and shows up the background. The disadvantage can to a certain extent be overcome by coating the record with a clear varnish as soon as it is removed from the recording instrument.

Some instruments utilise electrical markers, producing either a perforated trace by discharge through thin paper or a violet trace on iodinised paper.

Undoubtedly the clearest type of record is the photographic variety. Many recording vibrographs and other oscillation-measuring instruments are arranged so that the stylus is replaced by a moving light-spot, which records directly on to moving photo-

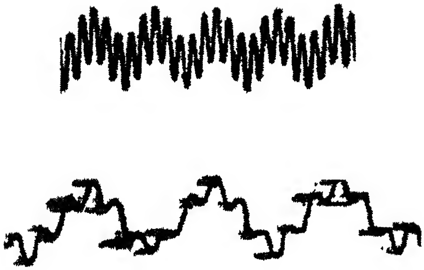


FIG. 1.—Blurred traces made by marking pens: (a) "fuzziness" caused by incorrect inking; (b) smudging due to inadequate arrangements for drying.

FIG. 2.—Multiple trace recorded photographically. See page 229 for explanation of timing marks. (de Havilland Aircraft Co.)

graphic film or paper. The light source is usually a fixed lamp, the excursion of the light-spot being obtained from the motion of an intervening mirror. This mirror may be actuated mechanically, or it may be mounted on the coil of a sensitive galvanometer, according to the nature of the recording system. Fig. 2 shows a multiple trace recorded directly on to photographic paper by twelve galvanometers ; the record also includes various timing indications.

Records on photographic paper are superior to those on film. The black trace on a white opaque ground is more definite, and imposes less eye-strain, than the opaque trace on the transparent ground. Moreover, envelopes and other construction lines can be marked easily with a pencil on the paper, while a scratch with a sharp-pointed stylus is required for film. If desired, an opaque print can always be made from a transparent "negative," but this procedure is uneconomical compared with the immediate production of an opaque record by using photographic paper instead of film ; furthermore, such a print would generally be "white-on-black," i.e. with a white trace on a black ground, and construction lines would have to be drawn in white ink. Most of the standard film sizes are available in paper form (35 mm. with or without sprocket holes, for example).

It may be noted that for high speeds of passage of the film through the camera it is advisable to use friction driving rollers, rather than sprockets, to avoid jamming.

3. Linearity of response.

Of paramount importance, where the analysis of a large number of records is concerned, is the linearity of the overall magnification of the recording system, or more shortly the "linear response." If the physical variation which is being measured is Y, and the excursion of the corresponding trace is y, then to a small increment δY in Y there corresponds a small increment δy in y, and the response is linear if

$$\frac{\delta y}{\delta Y} = M, \qquad (3.1)$$

where M is a constant, independent of the value of Y. In this case, integration of (3.1) gives

$$y = MY + C, \qquad (3.2)$$

where C is a constant of integration and refers to the choice of datum level. The equation (3.2) determines the *response curve* for the recording system, and this is in the form of a straight line, the slope of which affords a measure of the magnification.

If for any reason the response curve is not linear, *distortion* is produced. In Fig. 3 the vertical trace represents the physical variation and the horizontal trace represents the recorded waveform when the response curve has the non-linear form shown. If the distance between the lines A and B is a, and that between the lines C and D is b, the ratio a/b gives an indication of the distortion in the waveform shown in full lines. In the original variation

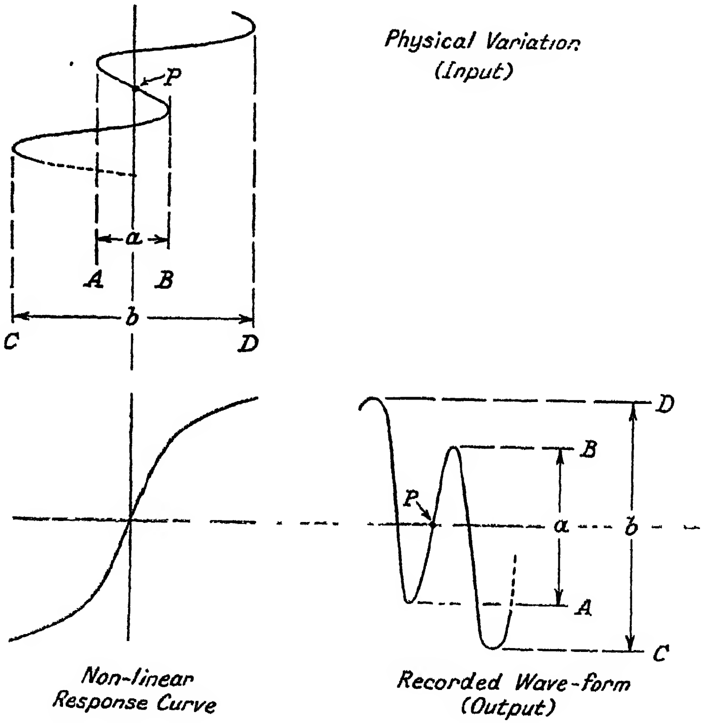


FIG. 3.—Effect of non-linear response of recording system. The vertical trace is the physical variation and the horizontal trace is the recorded waveform. (Compare Fig. 4).

$a/b = 0{\cdot}30$, and in the recorded waveform the ratio is $0{\cdot}64$, an increase of about 110 per cent.

The response curve shown in Fig. 3 is skew-symmetrical about the mean position, and the recorded waveform is still skew-symmetrical about P. Displacement of the mean position, or centre of skew-symmetry, can have a very marked effect on the distortion. In Fig. 4 is shown an extreme case, where the centre of skew-symmetry has been displaced a large distance from the position occupied in Fig. 3. The ratio a/b has the value $0{\cdot}20$, and the re-

corded waveform is no longer skew-symmetrical about P. On the other hand, it is evident that a condition of symmetry would be unaffected by this distortion.

From a study of Figs. 3 and 4 it will be evident that analysis of such waveforms cannot be performed without first converting the recorded values back to the true undistorted values. Such a process is excessively time consuming, and while it may be possible to employ this method in one or two instances, the source of the

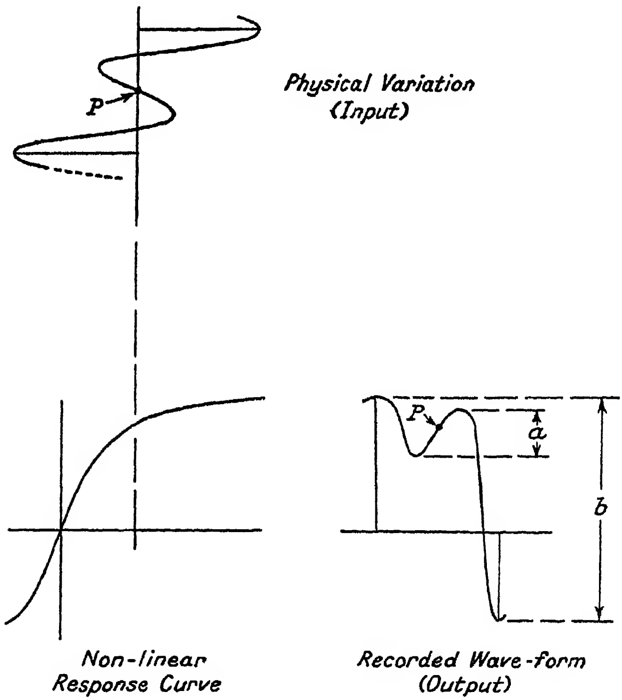


FIG. 4.—An extreme case of non-linearity of response.

distortion should be removed if it is required to analyse a large number of waveforms produced by a system with non-linear response.

4. Adequate length of record.

It cannot be too strongly emphasised that an adequate length of record is essential for accuracy in analysis. The estimation of what constitutes adequacy in this respect must take into account the fundamental frequencies of any variations likely to be present in the waveform, and also possible irregularities in the variation being recorded. Thus, for example, a torsiograph record taken

from an internal combustion engine (4-stroke) running on a brake bed has a fundamental wavelength corresponding to two revolutions of the engine, and in perfect operation the waveform will repeat accurately after successive intervals representing this period. At the same time, it must be borne in mind that perfect operation cannot be achieved, and there is likely to be some irregularity from cycle to cycle owing to uneven ignition or carburation in the engine. It is advisable, therefore, to take records extending over a number of cycles—say 10 cycles, or 20 revolutions—in order to observe the extent of this fluctuation and to estimate the average harmonic contents. There is no reason why the fluctuation should be periodic, and the best method of attack will be to analyse separate cycles of the main variation and to average the results.

The cycle of the main variation in a torsiograph or strain-gauge record taken from an aircraft engine/propeller combination, on the other hand, is likely to be very long if there is a reduction gear. This point has already been mentioned in the text (see pp. 47, 221). With a gear-ratio of 4 : 9 the cycle of the complete variation, including harmonics of both the engine and the propeller fundamental components, will extend over 18 engine revolutions or 8 propeller revolutions. It is therefore necessary in this case to record over at least 18 engine revolutions in order to obtain one complete cycle of the variation, and the effects of possible irregular fluctuations make it advisable to extend this range considerably.

In general, it is best to record at least 6 cycles of the complete variation, in order to estimate the average conditions. In the case of the gear-ratio 4 : 9 referred to above, this procedure would entail recording over about 120 engine revolutions. At a speed of 1800 R.P.M. the corresponding time would be 4 seconds, giving a length of record of about 18 feet with a film speed of 54 in. per second. With an average number of records to a test "run," taken at, say, 50 R.P.M. intervals over a speed range from 1500 to 3000 R.P.M., the complete film would be over 500 ft. long. Considerations of economy often require the ideal procedure to be modified.

In more general applications, the adequate length of record is governed by the accuracy required in the frequency determination. It is always advisable to have as long a record as other circumstances permit.

In connection with records obtained by photographing the screen of a cathode-ray tube, it will be seen that these considerations favour the adoption of continuous recording by means of a special camera, rather than taking single-shot photographs on plates or cut film. In continuous recording it is necessary to remove the

PLATE II

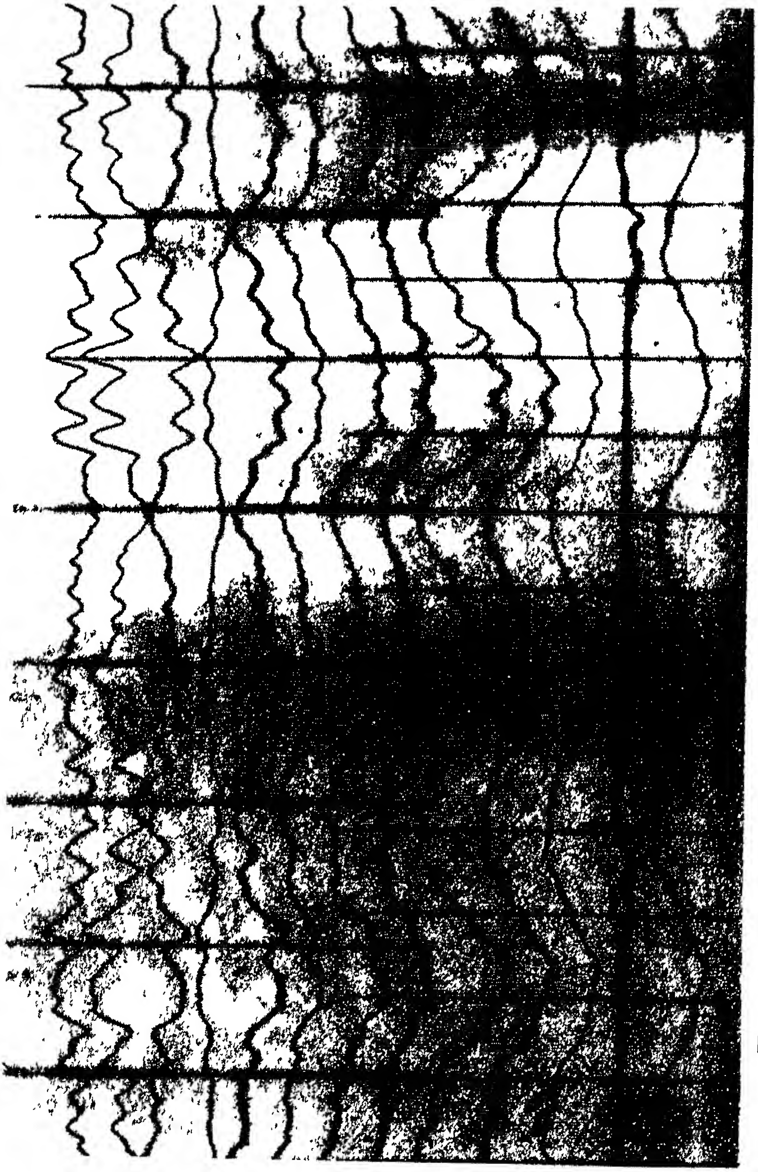


Fig. 5.—Multiple trace with three distinct time reference markings (see p. 229).
(de Havilland Aircraft Co.)

time-base "sweep" of the light-spot, so that the cathode-ray tube acts merely as a very sensitive galvanometer.

5. Time-reference markings.

With recording cameras it is usually very difficult to obtain a substantially constant film speed. It would of course be possible to run the camera continuously, switching-in the galvanometers when required, but such a procedure is very wasteful of film. The usual scheme is to stop the camera at the end of each record, re-starting at the beginning of the next record. With the usual motor arrangement and the normal length of record, a fair proportion of the running time is occupied in acceleration and deceleration, and even during the period when the camera is running at top speed there may be a slight surge. For this reason it is essential to provide sufficient time-reference markings. The record in Fig. 2 incorporates two such markings: first, the vertical lines indicate time-intervals of 1/100 second, and are produced by means of a synchronous motor running at A.C. mains frequency (3000 R.P.M.); the motor drives a shaft with a small diametral hole, through which light passes twice each revolution from a fixed lamp on to the record. Secondly, the second trace up from the bottom of the record indicates revolutions of the propeller, representative vibration stresses in which are recorded by the remaining traces. With this arrangement it is possible to obtain a very accurate check on the running speed of the propeller. Thus, in the original record from which the diagram is reproduced, the distance between consecutive revolution markings is 2·45 in., while the average film speed at this part of the record is 48·0 in. per second; hence the propeller speed is 48·0/2·45 = 19·6 R.P.S., or 1180 R.P.M., to the nearest 10 R.P.M.

In the record shown in Fig. 5, the upper set of vertical lines indicate successive half-revolutions of the engine, the other markings being as in Fig. 2. These engine-revolution indications are applied in a manner similar to that adopted for the time markings, by means of a synchronous motor operating on the output from a small generator coupled to the engine. In the original record the distance between consecutive propeller revolution markings is 4·40 in., and the average film speed at this part of the record is 54·9 in per second, so that the propeller speed is 54·9/4·40 = 12·5 R.P.S., or 750 R.P.M.; the length of record occupied by three engine revolutions (six timing intervals) is 5·88 in., so that the engine speed is 54·9 × 3/5·88 = 28·0 R.P.S., or 1680 R.P.M. The ratio 750/1680 = 0·446, and since the gear ratio was 4 : 9 the accuracy obtainable by using only a small portion of the record can be seen; naturally,

this accuracy can be improved by taking measurements over a longer portion of the record, and use of the accelerating and decelerating portions is not precluded

Fig 6 shows part of an accelerating record The paper moves from left to right, so that the right-hand end of the record refers to a time prior to that represented by the left-hand end. There is about 50 per cent increase in paper speed during the portion illustrated, and it will be seen that such a variation does not invalidate the envelope method of analysis or the determination of the frequencies of predominant components.

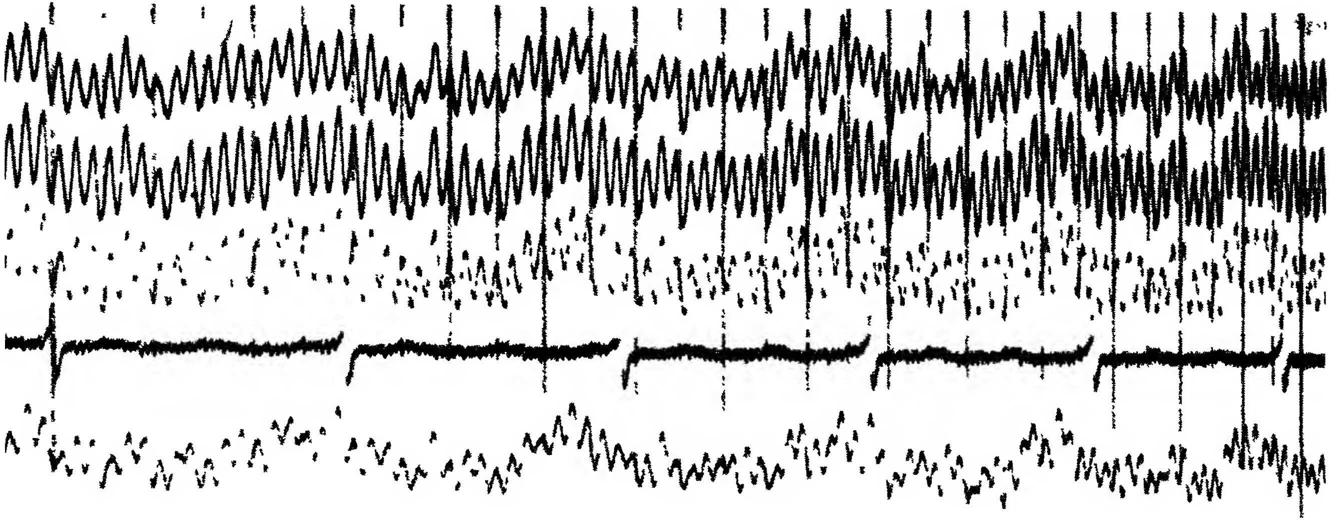


FIG. 6.—Accelerating record The paper speed at the left hand end (later) is about 50 per cent greater than at the right hand end (earlier) (de Havilland Aircraft Co)

[Facing page 230

CHAPTER X

LISSAJOU FIGURES

1. Introductory.

The waveforms considered hitherto in the text have been those generated by the variation of a single dependent variable. Another class of waveforms produced and studied in some fields of investigation are the Lissajou figures; these figures are produced by the motion of a point whose plane Cartesian co-ordinates both vary sinusoidally. The shape of the figure depends upon the ratio of the frequencies of the co-ordinate variations; when this is an integer, or can be expressed as the ratio of two small integers, the figure is fairly simple, but when it cannot be expressed in this form a very intricate pattern is obtained.

In particular, the application of sinusoidal voltage variations simultaneously to the two pairs of deflector plates in a cathode-ray tube produces the figures on the screen. Consequently, observation of a cathode-ray tube arranged in this manner provides a very convenient method of performing certain studies in connection with the behaviour of electric circuits: for example, the phase-distortion of an amplifier can be measured in this way, as there is a very simple relation between the shape of the Lissajou figure and the phase difference between the two injected sine-waves if these are of the same frequency.

Attention is here confined to the simpler figures [some very good illustrations of the more complicated variety are given by Osgood, reference 1 in the Bibliography at the end of the book]. The generation of the figures is first considered and illustrated; the determination of the frequencies of the co-ordinate variations from the form of the figure is then outlined, and the chapter concludes with a description of the method of determining the phase-difference between two sinusoidal waves of the same frequency, by measurement of the corresponding Lissajou figure.

2. Generation of Lissajou figures: equal frequencies.

Consider first the case of equal frequencies of variation of the two plane Cartesian co-ordinates which specify the position of the moving point Let these co-ordinates be x, y, and let them be expressible in terms of a common parametric variable t as

$$\left.\begin{aligned} x &= \sin(t+\phi) \\ y &= \sin t \end{aligned}\right\} \qquad (2.1)$$

Consideration is thus restricted to the case of equal amplitudes in the two variations. Expanding the first equation of (2.1),

$$x = \cos\phi \,.\, \sin t + \sin\phi \,.\, \cos t, \qquad (2.2)$$

and from the second equation of (2.1),

$$\left.\begin{aligned}\sin t &= y\\ \cos t &= \sqrt{1-y^2}\end{aligned}\right\} \qquad (2.3)$$

Substituting from (2.3) in (2.2),

$$x = y \ \cos\phi + \sqrt{1-y^2} \,.\, \sin\phi,$$
$$(1-y^2)\sin^2\phi = (x - y \,.\, \cos\phi)^2.$$

Rearranging this last equation,

$$x^2 + y^2 - 2xy \,.\, \cos\phi = \sin^2\phi. \qquad (2.4)$$

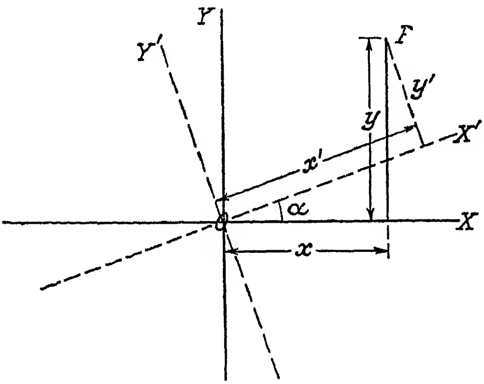


FIG. 1.—Rotation of reference axes.

In general, (2.4) represents an ellipse whose axes are inclined to the co-ordinate axes. Referring hereafter to the co-ordinate axes as the "injection axes," consider the effect of expressing (2.4) in terms of co-ordinates x', y', referred to new axes at an angle α to the injection axes (Fig. 1). Then

$$\left.\begin{aligned}x &= x'\cos\alpha - y'\sin\alpha\\ y &= x'\sin\alpha + y'\cos\alpha\end{aligned}\right\} \qquad (2.5)$$

Substituting these values in (2.4),

$$(x'\cos\alpha - y'\sin\alpha)^2 + (x'\sin\alpha + y'\cos\alpha)^2$$
$$- 2\,(x'\cos\alpha - y'\sin\sigma)(x'\sin\alpha + y'\cos\alpha)\cos\phi = \sin^2\phi,$$

i.e.

$$x'^2(1 - 2\sin\alpha \,.\, \cos\alpha \,.\, \cos\phi) + y'^2(1 + 2\sin\alpha \,.\, \cos\alpha \,.\, \cos\phi)$$
$$- 2x'y'(\cos^2\alpha - \sin^2\alpha)\cos\phi = \sin^2\phi. \qquad (2.6)$$

The new co-ordinate axes are principal axes of the ellipse if the term in xy vanishes, i.e. if

$$\cos^2\alpha = \sin^2\sigma,$$

or

$$\tan\alpha = \pm 1.$$

Taking $\alpha = 45°$, (2.6) reduces to

$$x'^2(1 - \cos\phi) + y'^2(1 + \cos\phi) = \sin^2\phi = 1 - \cos^2\phi,$$

or

$$\frac{x'^2}{1 + \cos\phi} + \frac{y'^2}{1 - \cos\phi} = 1. \qquad (2.7)$$

(2.7) is the equation of an ellipse referred to its principal axes, and comparing it with the standard equation to the ellipse, i.e. with

$$\frac{x'^2}{a^2} + \frac{y'^2}{b^2} = 1,$$

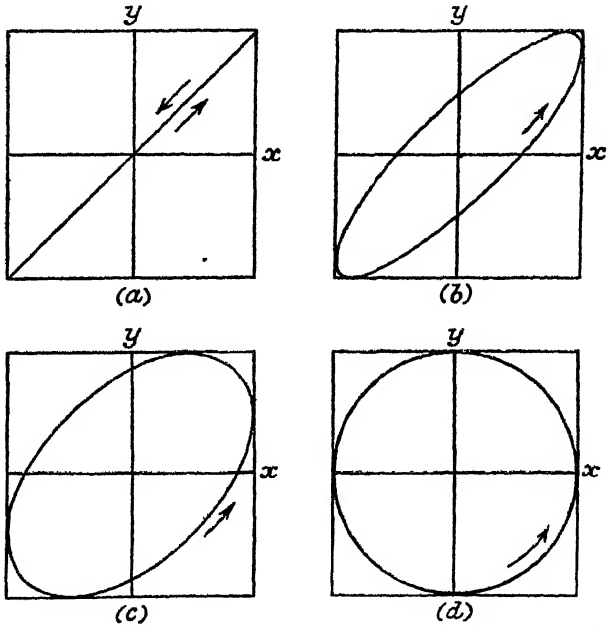


FIG. 2.—Simultaneous sinusoidal variations in two directions at right-angles, of the form (2 . 1) (equal frequencies). (*a*)–(*d*) represent the cases $\phi = 0$, 30°, 60° and 90° respectively.

it is seen that

$$\frac{b^2}{a^2} = \frac{1 - \cos\phi}{1 + \cos\phi} = \frac{2\sin^2\phi/2}{2\cos^2\phi/2},$$

whence

$$b/a = \tan\phi/2, \text{ numerically.} \qquad (2.8)$$

The Lissajou figure is therefore an ellipse, whose principal axes are inclined at an angle of 45° to the injection axes, and whose eccentricity depends upon the phase-difference ϕ. Some typical cases are illustrated in Fig. 2.

At (a) in the diagram the phase-difference ϕ is zero, and the figure is simply a portion of the line $y = x$. At (b) and (c) the cases $\phi = 30°$ and $60°$ are illustrated. The figure is an ellipse, and it can be seen that as ϕ is increased the ellipse becomes "thicker," i.e. the ratio b/a increases. At (d), $\phi = 90°$ and the ellipse degenerates into a circle; equation (2.8) gives $b/a = 1$ for this case, as would be expected. The arrows on the diagrams show the direction in which the curves are traced by the moving point as the parameter t is increased.

If ϕ is greater than $90°$, the equation (2.8) gives b as greater than a, so that the y' axis becomes the major axis. Let

$$\phi = 180° - \psi.$$

Then $\cos \phi = - \cos \psi$, and $\sin^2 \phi = \sin^2 \psi$, so that (2.4) becomes

$$x^2 + y^2 + 2xy \,.\, \cos \psi = \sin^2 \psi. \qquad (2.9)$$

The curve represented by (2.9) can be obtained from that represented by (2.4) by substituting ψ for ϕ and changing the sign either of x or of y—i.e. by taking the mirror-image of the curve either in the x-axis or in the y-axis.

Since $\cos (360° - \phi) = \cos \phi$, and $\sin^2 (360° - \phi) = \sin^2 \phi$, the equation (2.4) is unaltered if $(360° - \phi)$ is substituted for ϕ. The figure for $\phi = 210°$ will therefore be identical with that for $\phi = 150°$, and so on. Intuition suggests, however, that there must be *some* difference between the two diagrams. To determine wherein lies this difference, reference must be made to the original parametric equations (2.1). Let $\phi = 360° - \psi'$, then these equations become

$$\left.\begin{array}{l} x = \sin (t - \psi') \\ y = \sin t \end{array}\right\} \qquad (2.10)$$

Rearranging,

$$\left.\begin{array}{l} - x = \sin (- t + \psi') \\ - y = \sin (- t) \end{array}\right\} \qquad (2.11)$$

Now, changing the sign of both x and y in equation (2.4), which is the parametric equation to the ellipse, effects no alteration in the curve; so that the equations (2.11) represent the same curve as

$$\left.\begin{array}{l} x = \sin (- t + \psi') \\ y = \sin (- t) \end{array}\right\} \qquad (2.12)$$

Thus the curve for a phase-difference $\psi' = 360° - \phi$ is the same as that for a phase-difference ϕ but is traced by the moving point in the opposite direction; Fig. 2 *a-d* can be taken to represent the cases $\phi = 270°$ to $360°$ if the arrows are reversed.

3. Other simple frequency ratios.

When the variations in x and y have different frequencies, more complicated figures are obtained. The standard form for the parametric equations will at first be taken as

$$\left.\begin{aligned} x &= \sin(nt+\phi) \\ y &= \sin t \end{aligned}\right\} \quad . \quad . \quad . \quad (3.1)$$

where the amplitudes of the two variations are both unity.

Consider first the case $n = 2$. Then

$$x = \sin 2t \,.\, \cos\phi + \cos 2t \,.\, \sin\phi, \quad . \quad . \quad (3.2)$$

and $$\sin t = y,\ \cos t = \sqrt{1-y^2}.$$

Substituting for $\sin t$ and $\cos t$ in (3.2),

$$x = 2y\sqrt{1-y^2} \,.\, \cos\phi + (1-2y^2)\sin\phi,$$

or $$4y^2(1-y^2)\cos^2\phi = [x + (2y^2-1)\sin\phi]^2.$$

The intrinsic equation to the curve therefore becomes

$$4y^2(1-y^2) = x^2 + 2x(2y^2-1)\sin\phi + \sin^2\phi. \quad . \quad (3.3)$$

Several typical cases are illustrated in Fig. 3. First, if $\phi = 0$, equation (3.3) reduces to

$$\left.\begin{aligned} 4y^2(1-y^2) &= x^2 \\ \text{or} \qquad 4y^4 - 4y^2 + x^2 &= 0 \end{aligned}\right\} \quad . \quad . \quad . \quad (3.4)$$

For any value of y, not greater numerically than unity, there are two values of x; and the second equation of (3.4) gives

$$y^2 = \tfrac{1}{2}(1 \pm \sqrt{1-x^2}),$$

so that there are in general four values of y for any value of x not greater numerically than unity. This case is illustrated at (*a*), Fig. 3. (*b*) and (*c*) in the same diagram illustrate the cases $\phi = 30°$ and $\phi = 60°$, the curves for which are of much the same general form as for $\phi = 0$.

When $\phi = 90°$, (3.3) reduces to

$$4y^2(1-y^2) = x^2 + 2x(2y^2-1) + 1.$$

This equation can be rewritten as

$$(2y^2)^2 + x^2 + 1^2 + 2(2y^2)x - 2(2y^2)1 - 2x \,.\, 1 = 0,$$

i.e. $$(2y^2 + x - 1)^2 = 0,$$

or $$x = 1 - 2y^2. \quad . \quad . \quad . \quad . \quad (3.5)$$

[This result can also be obtained at once from equations (3.1), for $(x = \cos 2t = 1 - 2\sin^2 t = 1 - 2y^2]$.

In this case the curve is part of a parabola, as shown at (*d*) in the diagram.

For values of ϕ greater than 90°, but not greater than 180°, let $\phi = 180° - \psi$. Then equation (3.3) remains unaltered, except that ψ is substituted for ϕ, since $\sin(180° - \psi) = \sin\psi$. Hence the diagrams *a-c* in Fig. 3 serve also to illustrate the cases $\phi = 180°$, 150° and 120°, respectively; as with the frequency ratio 1 : 1, the curves are traced in the reverse direction for ϕ greater than 90°.

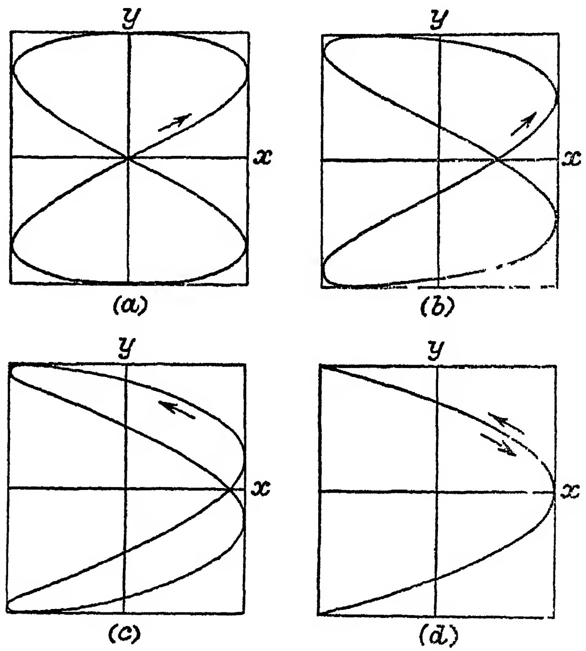


Fig. 3.—Frequency-ratio 2 : 1. See text for details of the phase relationships in the various cases.

For values of ϕ greater than 180°, let $\phi = 360° - \psi'$. Then equation (3.3) becomes

$$4y^2(1 - y^2) = x^2 - 2x(2y^2 - 1)\sin\psi' + \sin^2\psi'. \quad . \quad (3.6)$$

Now since in (3.6) the only term involving an odd power of x is the negative of the corresponding term in (3.3) all the other terms remaining unaltered (except that ψ' is substituted for ϕ), it follows that (3.6) could be obtained from (3.3) by changing the sign of x; thus the curves for ϕ greater than 180° are the mirror-images in the y-axis of the curves for the corresponding values of $(360° - \phi)$.

Fig. 4 depicts typical cases of the form

$$x = \sin(3t + \phi),$$
$$y = \sin t.$$

The diagrams *a-d* correspond to the values $\phi = 0$, $30°$, $60°$ and $90°$. With the less simple values of the frequency ratio the derivation of the intrinsic equation becomes more difficult, and the general properties are more conveniently derived from the parametric equations :

$$\left.\begin{aligned} x &= \sin(3t + \phi) \\ y &= \sin t \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (3.7)$$

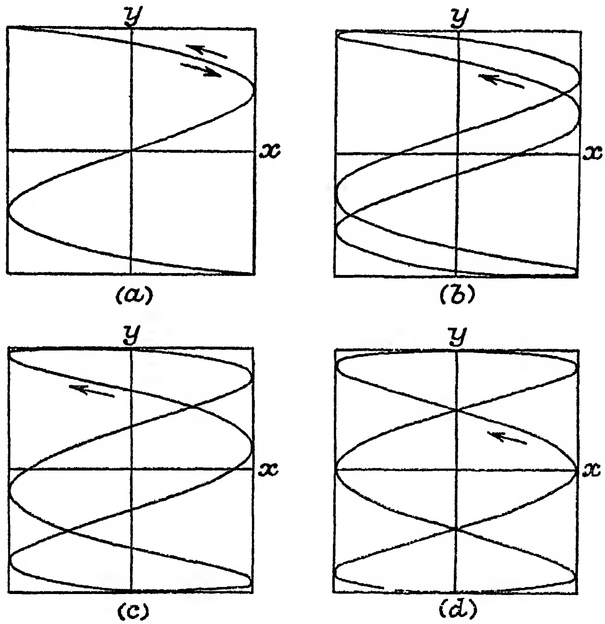


FIG. 4.—Frequency-ratio 3 : 1. See text for details of the phase-relationships in the various cases.

For values of ϕ greater than 90° but not greater than 180°, let $\phi = 180° - \psi$. Equations (3.7) become

$$\left.\begin{aligned} x &= -\sin(3t - \psi) \\ y &= \sin t \end{aligned}\right\},$$

or

$$\left.\begin{aligned} x &= \sin(-3t + \psi) \\ y &= -\sin(-t) \end{aligned}\right\} \quad . \quad . \quad . \quad (3.8)$$

The equations (3.8) represent the same curve as do

$$\left.\begin{aligned} x &= \sin(3t + \psi) \\ y &= -\sin t \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (3.9)$$

except that the curve is traced in the reverse direction as t is increased; and the form of (3.9) shows that the curves for ϕ greater than 90° are the mirror-images in the x-axis of the curves for the corresponding values of $(180° - \phi)$.

For values of ϕ greater than 180°, let $\phi = 360° - \psi'$. Then equations (3.7) become

$$\left.\begin{aligned} x &= \sin(3t - \psi') \\ y &= \sin t \end{aligned}\right\}$$

i.e.

$$\left.\begin{aligned} x &= -\sin(-3t + \psi') \\ y &= -\sin(-t) \end{aligned}\right\} \quad . \quad . \quad (3.10)$$

The form of these equations shows that the curves for values of ϕ greater than 180° are obtained from those for the corresponding values of $(360° - \phi)$ by changing the sign of both x and y—i.e. by rotating the diagram in its own plane through an angle of 180°. It should be noted that such a rotation effects no alteration in this case $(n = 3)$.

Consider now the more general form

$$\left.\begin{aligned} x &= \sin(nt + \phi) \\ y &= \sin mt \end{aligned}\right\} \quad . \quad . \quad . \quad (3.11)$$

It will be seen that the facts concerning the curves for ϕ greater than 90° are as follows:

(i) If the phase-angle ϕ is greater than 90° but not greater than 180°, the curve is the mirror-image in the x-axis of the curve corresponding to the phase-angle $(180° - \phi)$.

(ii) If the phase-angle ϕ is greater than 180°, the curve is obtained from that for the phase-angle $(360° - \phi)$ by rotating it in its own plane through an angle of 180°.

The proof of these statements follows exactly the same lines as does that given above for the case $n = 3$, $m = 1$.

The fact that these rules are not exactly the same in form as those given above for some of the simpler cases results from inherent symmetry or skew-symmetry in the curves in those particular examples; it will be found that the rules give the correct result in all cases.

Figs. 5 and 6 illustrate some typical cases of the forms

$$\left.\begin{aligned} x &= \sin(3t + \phi) \\ y &= \sin 2t \end{aligned}\right\} \quad . \quad . \quad . \quad (3.12)$$

and

$$\left.\begin{aligned} x &= \sin(5t + \phi) \\ y &= \sin 2t \end{aligned}\right\} \quad . \quad . \quad . \quad (3.13)$$

respectively.

Note.—In all the cases considered in the preceding sections the amplitudes of both the co-ordinate variations have been taken as unity. The effect of altering these amplitudes is merely to change

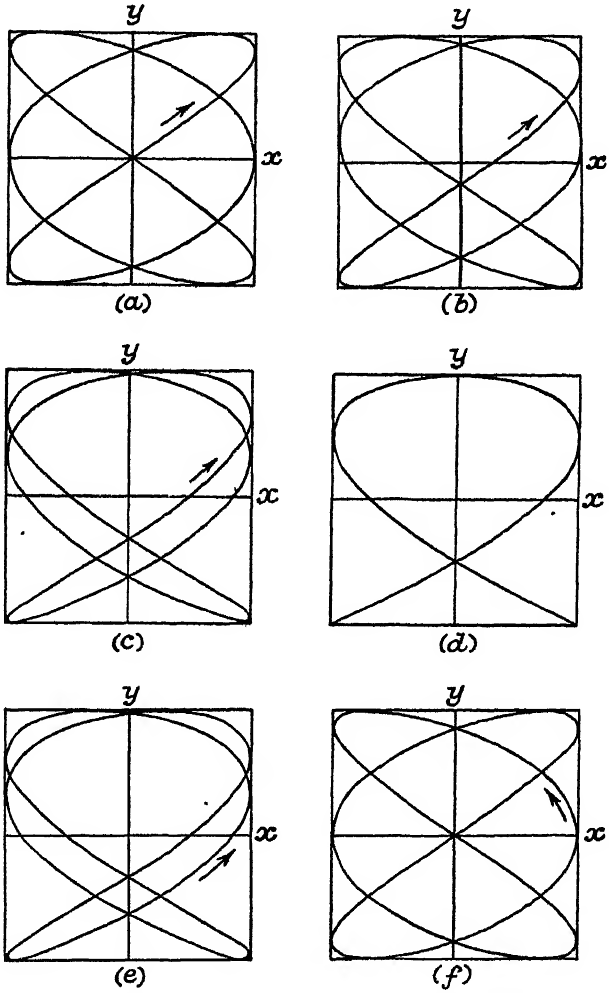


FIG. 5.—Frequency-ratio 3 : 2. (*a*)–(*f*) represent the cases $\phi = 0$, 30°, 60°, 90°, 120° and 180° respectively.

the relative proportions of the figures, without altering their general form. Thus, for example, with equal frequencies and $\phi = 90°$ the figure is an ellipse instead of a circle if the amplitudes are unequal, the ellipse having the injection axes as its principal axes.

4. Determination of frequency ratio.

The ratio of the frequencies of the co-ordinate variations can easily be determined from a given figure. Starting from any convenient point in the curve, the *complete* curve is traced out with a pencil until the starting point is reached again with the same direction of motion as at first; during this operation a note is made of the number of complete oscillations made from side to side. The process is repeated, the number of complete oscillations up-and-down being noted. The ratio of these two numbers is the required frequency ratio.

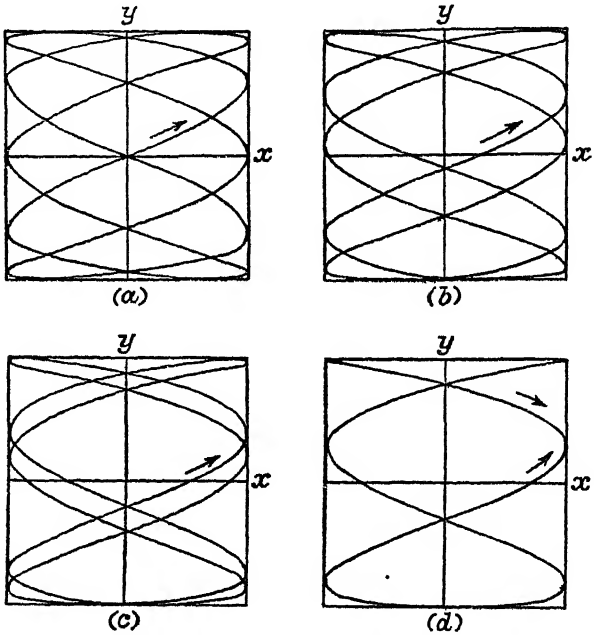


FIG. 6.—Frequency-ratio 5 : 2.

Thus, for example, starting from the point A in Fig. 7*a* and tracing out the curve, a cycle of horizontal motion occurs in the portion from A to B, another cycle from B to C, and half a cycle from C to D; retracing the curve from D back to A, a total of five cycles of horizontal motion is obtained. Similarly, the portion from A to D comprises one cycle of vertical motion, so that the complete circuit extends over two cycles of vertical motion. The frequency ratio is therefore 5 : 2.

The more complicated pattern in Fig. 7*b* is amenable to the same treatment. Thus, cycles of horizontal motion occur in the

portions AB, BC, CD, DE, EF, FG and GA, while cycles of vertical motion occur in the portions AH, HJ, JK, and KA. The frequency ratio is therefore 7 : 4.

Another method of determining the frequency ratio is as follows: draw a line parallel to the x-axis which does not pass through any intersections of different parts of the curve, and count the number of points in which the curve intersects the line; repeat this procedure with a line drawn parallel to the y-axis (the lines must *not* be drawn so as to *touch* the exterior lobes of the curve—i.e. the lines must not be part of the bounding rectangle). The ratio of the two numbers so found will be the required frequency ratio: the number of intersections on the line parallel to the x-axis is proportional to the frequency of the y-variation, and *vice versa*. This rule is applicable to all cases, whether the curve is a closed one or consists of a " doubled " line, as in Fig. 5*d*.

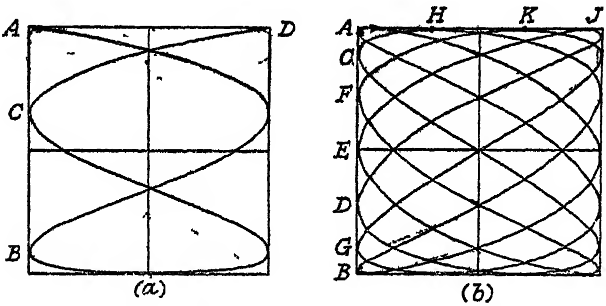


FIG. 7.—Determination of frequency-ratio.

5. Determination of phase-difference.

The study of the Lissajou figure resulting from the orthogonal composition of two sinusoidal variations of the same frequency affords a very convenient method of measuring the phase-difference between the two variations. The determination is easiest when the amplitudes of the two variations are equal, for then formula (2.8) on page 233 gives the value of ϕ in a very simple form:

$$\phi = 2 \tan^{-1}\frac{b}{a}, \quad . \quad . \quad . \quad . \quad (5.1)$$

where a and b are the semi-intercepts on axes at angles 45° and 135° respectively, with the x-axis, and ϕ is the phase-angle in the general form (2.1), page 231. Thus where the major axis is inclined at an angle of 45° with the x-axis, b/a is the ratio of the minor axis to the major axis, but when the angle between the major axis and the x-axis is 135°, b/a is the ratio of the major axis to the minor axis.

(The angle referred to is that made in a positive sense, i.e. counter-clockwise, with the positive direction of the x-axis.)

Owing to the simplicity of the form of equation (5.1) it is advisable always to adjust matters so that the amplitudes of the co-ordinate variations are equal. This can easily be done with a cathode-ray tube by incorporating variable attenuators in one or both pairs of injection leads. If a pair of cross-lines is marked on the screen, the lines being of equal length, bisecting each other and lying along the injection axes, the procedure can be carried out as follows: switch off one co-ordinate variation, so that the trace on the screen becomes a straight line lying in one of the injection axes; adjust the input attenuator so that the amplitude of the variation parallel to this axis is the required value, the line just covering the marked line; repeat the process with the other co-ordinate variation, then switch both variations in and measure the major and minor axes of the ellipse.

If it should not be possible to make these adjustments, so that the amplitudes are not equal, the formula for calculating the phase-angle becomes slightly more complicated. If R is the ratio of the amplitude of the y-variation to the x-variation, the ratio b/a is given by

$$\frac{b^2}{a^2} = \frac{[(1 - \mathrm{R}^2)^2 + 4\mathrm{R}^2 \cos^2 \phi]^{\frac{1}{2}} . (1 + \mathrm{R}^2) - (1 - \mathrm{R}^2)^2 - 4\mathrm{R}^2 \cos^2 \phi}{[(1 - \mathrm{R}^2)^2 + 4\mathrm{R}^2 \cos^2 \phi]^{\frac{1}{2}} . (1 + \mathrm{R}^2) + (1 - \mathrm{R}^2)^2 + 4\mathrm{R}^2 \cos^2 \phi}.$$

This formula is established in a paper by Dr. M. K. B. Richards and the author, see reference 2 in the Bibliography; the rather cumbersome expression can be simplified * into:

$$\left.\begin{aligned} \sin \phi &= \frac{r(1 + \mathrm{R}^2)}{\mathrm{R}(1 + r^2)} \\ \text{where} \qquad r &= b/a. \end{aligned}\right\} \quad . \quad . \quad . \quad (5.2)$$

In cases where the physical nature of the problem does not indicate which of the two values of ϕ given by (5·2) is the correct one to take, it is preferable to employ the method described in the next paragraph.

Since the publication of the paper referred to, Mr. P. V. MacKinson has pointed out a much simpler method of determining the phase-angle in cases where the amplitudes are unequal. The lines $y = \mathrm{B}x/\mathrm{A}$ and $y = -\mathrm{B}x/\mathrm{A}$ intersect the ellipse at points P and Q (see Fig. 8) such that

$$\tan \phi/2 = \mathrm{OQ}/\mathrm{OP},$$

* Simplification due to Mr. C. E. A. Burnham.

so that equation (5.1) can be replaced by

$$\phi = 2 \tan^{-1} (OQ/OP). \quad . \quad . \quad . \quad (5.3)$$

This property can be established from the corresponding property for equal amplitudes by orthogonal projection, since the diagonals of the circumscribing square project into the diagonals of the circumscribing rectangle, and since they are equally inclined to the X-axis the intercepts cut off by the ellipse remain in the same ratio. The result can also be established by analysis as follows:

The parametric equations to the ellipse are

$$\left.\begin{aligned} x &= A \sin (t + \phi) \\ y &= B \sin t \end{aligned}\right\} \quad . \quad (5.4)$$

From these equations the intrinsic equation is derived, by eliminating t, as

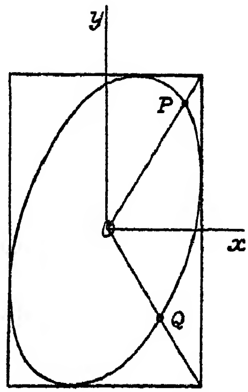


FIG. 8.—Alternate method of determining phase-angle in the case of equal frequencies.

$$x^2 - 2xy\frac{A}{B} \cos \phi + \frac{A^2}{B^2} y^2 = A^2 \sin^2 \phi.$$

P is the point of intersection of the ellipse and the line $y = Bx/A$, and is given by the co-ordinates

$$x = A \cos \phi/2,$$
$$y = B \cos \phi/2.$$

Hence $$OP = \sqrt{A^2 + B^2} \cos \phi/2.$$

Similarly, $$OQ = \sqrt{A^2 + B^2} \sin \phi/2,$$

so that $$\tan \phi/2 = OQ/OP.$$

It should be noted that in any case, only the phase-difference is determined by this method, and no information can be gained as to which variation "leads" the other. The distinction between the "leading" and "lagging" variations can usually be made by a consideration of the physical characteristics of the system under test; or, alternatively, when a cathode-ray tube is used, the direction of travel of the "spot" round the figure can be determined by special means as described in reference 3 (Bibliography, p. 268).

APPENDIX I

CHOICE OF SINE-WAVES AS BASIC COMPONENTS OF WAVEFORMS

ONE aspect of waveform analysis has received only passing mention in the body of the text, namely the reasons why sine-waves are chosen as the basic components of waveforms. There are five main reasons for this choice, and these will be discussed in turn.

In general the term "sine-wave" or "sine function" is to be understood as referring to the form $a \,.\, \sin (x + \phi)$.

1. Natural occurrence of sine functions.

The generation of a sine-wave by projection from a uniformly rotating vector (Chapter I, p. 8) is a phenomenon which occurs in many machines. For example, the centrifugal force vector acting through the mass centre of an unbalanced rotor revolving with uniform angular velocity can be resolved in two directions mutually at right angles, the projections having sinusoidal variations. Similarly, the electric voltage generated when a single turn of wire is rotated uniformly inside a uniform magnetic field is a sine function of the angular position of the rotor.

The free vibration of a simple mechanical system having one degree of freedom is likewise sinusoidal, in the absence of damping forces (see reference 1 in the Bibliography). The sine function is inherent in the solution of the differential equations expressing the motion of mechanical vibrating systems, as also in the solution of the equations expressing the propagation of electro-magnetic waves. Thus the sine function appears wherever the existence of vibrations is investigated mathematically.

2. Forced oscillation.

In the study of forced oscillations, whether mechanical or electrical, it is found that the response to a sinusoidal force or voltage is a sinusoidal displacement or current of the same frequency; this result depends ultimately upon the fact that all the derivatives of the sine function are themselves sine functions.

3. "Smoothness."

The sine function is the simplest periodic function which fulfils the condition that itself and all its derivatives should be continuous. The mathematical condition of continuity in a function and its derivatives corresponds to the intuitive concept of "smoothness." The functions whose graphs are shown in Figs. 5, 8 and 7 of Chapter VI, considered in that order, constitute a series of periodic functions in which each member is smoother than the preceding one. The square-peaked waveform in Fig. 5 is itself discontinuous; in Fig. 8 the graph is continuous but the first derivative of the function (the slope of the graph) is discontinuous, and in Fig. 7 the graph and its first derivative are continuous but the second derivative is discontinuous. There are, it is

true, many functions which are periodic and fulfil the required condition, but the sine function is the simplest Every derivative of the sine-wave is a sine-wave, with a suitable phase-relationship to the original, and thus every derivative is continuous.

4. Simplicity.

It is frequently contended that the sine-wave is the "simplest" periodic variation. Even if consideration is limited to continuous variations, such a contention is questionable; it is doubtful whether the sine-wave is essentially simpler than the saw-tooth waveform shown in Fig. 8c of Chapter VI, page 161; certainly the saw-tooth wave is simpler to draw than the sine-wave. The product of two sine functions can, however, be expressed as the sum or difference of two related sine functions, and the sum of two sine-functions of the same period can be expressed as a single sine function of the same period, so that there is much justification for the contention that the sine is the simplest periodic function as regards ease of manipulation. It is also the first periodic function to be encountered in the normal sequence of education, and for that reason is very familiar to students.

On account of the almost universal application of the trigonometric functions to all classes of calculations, tables of the sine and cosine are widely circulated, and any combination of sine functions is easily constructed.

5. Orthogonal functions, Fourier analysis.

There exists a whole class of mathematical series of functions which are said to be orthogonal and normal for a given range of the basic variable. If such a series of functions is $f_1(x)$, $f_2(x)$, $f_3(x)$, etc., and the given range is from $x = a$ to $x = b$, the condition of orthogonality is that

$$\int_a^b f_i(x) \,.\, f_j(x)\, dx = 0 \quad \text{if} \quad i \neq j \qquad . \quad . \quad . \quad (5.1)$$

and the condition of normality is that

$$\int_a^b [f_i(x)]^2\, dx = \text{K, where K is independent of } i. \qquad . \quad (5.2)$$

It is a well-known result in the theory of mathematical analysis that from any series of continuous functions which fulfil a certain condition (linear independence) it is possible to construct a series of orthogonal normal functions (see reference 2 in the Bibliography, from which place the first part of the following example is taken). For example, from the series of functions $1, x, x^2, x^3$, etc., the series

$$1,\ x,\ x^2 - 1/3,\ x^3 - 3x/5,\ x - 6x^2/7 + 3/35, \text{ etc.,} \qquad . \quad (5.3)$$

can be constructed, and is orthogonal for the range of x from -1, to 1; thus

$$\int_{-1}^{1} 1 \,.\, x\, dx = \Big[x^2/2\Big]_{-1}^{1} = 0,$$

$$\int_{-1}^{1} 1(x^2 - 1/3)dx = \left[\frac{x^3}{3} - \frac{x}{3}\right]_{-1}^{1} = 0,$$

$$\int_{-1}^{1} x(x^2 - 1/3)dx = \left[\frac{x^4}{4} - \frac{x^2}{6}\right]_{-1}^{1} = 0, \text{ etc.}$$

On the other hand,

$$\int_{-1}^{1} (1)^2\,dx = 2,$$

$$\int_{-1}^{1} (x)^2\,dx = 2/3,$$

$$\int_{-1}^{1} (x^2 - 1/3)^2\,dx = 8/45,$$

so that the series (5.3) is not normal in its present form. The series

$$1,\ \sqrt{3}x,\ \tfrac{3}{2}\sqrt{5}(x^2 - 1/3),\ \text{etc.}, \quad . \quad . \quad . \qquad (5.4)$$

is, however, normal: the value of K in the integral (5.2) is 2 for this particular series.

From a given set of periodic continuous functions a set of normal orthogonal periodic functions can be constructed in a similar manner, and any given arbitrary periodic function can then be expanded as the sum of a number of the derived orthogonal functions. The Fourier series of sine and cosine functions is a particular example of such an expansion.

Thus, for example, any periodic force imposed on a structure can be regarded as the sum of a number of sinusoidal forces at different frequencies, and the response of the structure is the sum of the (sinusoidal) responses to these various components (see reference 3 in the Bibliography).

APPENDIX II

UNIQUENESS OF FOURIER SERIES EXPANSION

THE proposition that the Fourier series expansion of an arbitrary periodic function is unique has been assumed as a truth in the text The rigorous proof of this proposition, known as Riemann's Theorem, is given in reference 4 in the Bibliography, and is far too lengthy and involved to be quoted here; neither is such a quotation necessary. It is possible, however, to outline a demonstration of the probable truth of the proposition.

Informal demonstration of Riemann's Theorem.

Let the function $y = f(x)$ be represented approximately by the Fourier series

$$\left.\begin{aligned} S(x) &= A_0 + \sum_{k=1}^{n} a_k \cos kx + \sum_{k=1}^{n} b_k \sin kx \\ &\text{and also by the series} \\ S'(x) &= A'_0 + \sum_{k=1}^{n} a'_k \cos kx + \sum_{k=1}^{n} b'_k \sin kx \end{aligned}\right\} \quad . \quad . \quad (1)$$

where k is integral These series terminate at the nth harmonic, so that they may be made to represent the function y at $(2n + 1)$ distinct values of x, say at $x_1, x_2, \ldots x_{2n+1}$ Then

$$\left.\begin{aligned} f(x_i) &= A_0 + \sum_{k} a_k \cos kx_i + \sum_{k} b_k \sin kx_i \\ \text{and} \quad f(x_i) &= A'_0 + \sum_{k} a'_k \cos kx_i + \sum_{k} b'_k \sin kx_i \end{aligned}\right\} \quad . \quad . \quad (2)$$

where i takes the values $1, 2, \ldots 2n + 1$ successively.

Two sets of $(2n + 1)$ linear equations are thus obtained; in each set there are $(2n + 1)$ unknown quantities (A_0, a_k and b_k; A'_0, a'_k and b'_k) and the coefficients $\cos kx_i$ and $\sin kx_i$ are the same in both sets. The equations (2) therefore have the same solutions in the unknown quantities, given by

$$\left.\begin{aligned} A_0 = A'_0 &= \frac{1}{2n+1} \sum_{i=1}^{2n+1} f(x_i) \\ a_k = a'_k &= \frac{2}{2n+1} \sum_{i=1}^{2n+1} f(x_i) \cos kx_i \\ b_k = b'_k &= \frac{2}{2n+1} \sum_{i=1}^{2n+1} f(x_i) \sin kx_i \end{aligned}\right\} \quad . \quad . \quad . \quad (3)$$

The arbitrary periodic function $y = f(x)$ can therefore be represented only by one harmonic series containing harmonics up to the nth which is accurate for the $(2n + 1)$ points $x_1, x_2, \ldots x_{2n+1}$, since $A_0 = A'_0$, $a_k = a'_k$, and $b_k = b'_k$.

If, now, n is increased but remains finite, the same result is found to be true, since the equations (3) are not affected by any alteration in n. As n is increased, the number $(2n + 1)$ of points at which the function is accurately represented by the series is also increased. If now n is increased without limit, it is probably true that the result still holds good, and the probable truth of the proposition has therefore been demonstrated.

APPENDIX III

EFFECT OF NON-HARMONIC COMPONENTS, OR OF CHOOSING A FALSE CYCLE

IF a periodic waveform is subjected to Fourier analysis, by mathematical or mechanical methods, it is essential to take a true cycle of the waveform as the data for the analysis. This Appendix describes briefly the effects of taking a false cycle, or of neglecting to observe the presence of a non-harmonic component superimposed on the main variation.

1. False analysis of non-harmonic component.

Let the period of the main variation be 2π radians in x. If p is not an integer, the wave

$$P = A_p \sin(px + \phi) = a_p \cos px + b_p \sin px \quad . \quad . \quad (1)$$

is non-harmonic with respect to the cycle of the main variation. The errors in the Fourier coefficients due to the presence of P will be :

$$\left.\begin{aligned} A'_0 &= \frac{1}{2\pi}\int_0^{2\pi} P \, . \, dx \\ a'_k &= \frac{1}{\pi}\int_0^{2\pi} P \, . \cos kx \, . dx \\ b'_k &= \frac{1}{\pi}\int_0^{2\pi} P \, . \sin kx \, . \, dx \end{aligned}\right\} \quad . \quad . \quad . \quad (2)$$

Substitution of (1) in (2) yields the results :

$$\left.\begin{aligned} A'_0 &= \frac{1}{2p\pi}[a_p \sin 2p\pi + b_p(1 - \cos 2p\pi)] \\ a'_k &= \frac{p}{\pi(p^2 - k^2)}[a_p \sin 2p\pi + b_p(1 - \cos 2p\pi)] \\ b'_k &= \frac{k}{\pi(p^2 - k^2)}[- a_p(1 - \cos 2p\pi) + b_p \sin 2p\pi] \end{aligned}\right\} \quad . \quad (3)$$

Example (i)

Suppose a periodic variation contains the terms

$$y = 0{\cdot}2 + \sin 4z + \cos 8z + 0{\cdot}5 \sin 9z \quad . \quad . \quad . \quad (4)$$

so that the harmonic coefficients are as given in Table I.
If the 9th harmonic is disregarded, the modified series S is .

$$S = 0{\cdot}2 + \sin 4z + \cos 8z$$

or $\quad S = 0{\cdot}2 + \sin x + \cos 2x$
where $\quad x = 4z \quad . \quad . \quad . \quad (5)$

The 9th harmonic of (4) can be put in the form

$$P = 0{\cdot}5 \sin 2{\cdot}25x \quad . \quad . \quad . \quad . \quad (6)$$

TABLE I

k	a_k ($A_0 = \frac{1}{2}a_0$)	b_k	A_k
0	0·4	—	0·2
4		1	1
8	1		1
9		0·5	0·5

and accordingly, if the waveform is analysed over the cycle 0 — 2π in x, the errors in the Fourier coefficients as given by (3) are as set forth in Table II.

TABLE II

k	a'_k ($A'_0 = \frac{1}{2}a'_0$)	b'_k
0	0·070	—
1	0·088	0·039
2	0·337	0·299
3	— 0·091	— 0·121
4	— 0·033	— 0·058

The Fourier coefficients for the waveform (4), analysed in this fashion, would therefore be found to be as set forth in Table III. The first column gives the harmonic reference numbers in terms of the cycle 0 — 2π in x, while the second column gives them in terms of the cycle 0 — 2π in z.

TABLE III

k (x)	k (z)	a_k ($A_0 = \frac{1}{2}a_0$)	b_k	A_k	Error in A_k
0	0	0·470	—	0·235	0·035
1	4	0·088	1·039	1·043	0·043
2	8	1·337	0·299	1·370	0·370
3	12	— 0·091	— 0·121	0·151	0·151
4	16	— 0·033	— 0·058	0·067	0·067

The last column but one in the table gives the values of the harmonic amplitudes as found by the analysis, and the last column gives the errors, which are the differences between these values and those given in Table I.

It will be observed that the errors are greatest in the harmonics which are nearest in frequency to the non-harmonic component.

The above example is a simple case. A further example is included, this being one of greater practical probability.

Example (ii)

The use of optical and mechanical analysers (pp. 212 and 219) has been suggested for the analysis of strain-gauge records taken from an aircraft propeller. It was at first thought that it would not be necessary to take a full cycle of the complex strain variation, containing both engine and propeller harmonics (p. 47), but only to take, in succession, a convenient number of cycles of the engine components and similarly for the propeller components.

Table IV gives the results obtained by analysis of the $3\frac{1}{2}$ E order waveform, i.e. the 7th harmonic of the 4-stroke engine cycle, the analysis being performed over one propeller revolution. The reduction-gear ratio is 9 : 4, so that the $3\frac{1}{2}$ E wave corresponds to a value $7\frac{7}{8}$ for p in (1). The cycle has been taken in such a manner that the non-harmonic component is a true sine-wave of unit amplitude, i.e. $a_p = 0$, $b_p = 1$. It will be noted that the error is greatest in the 8th harmonic, whose frequency is nearest that of the non-harmonic component.

TABLE IV

k	a'_k ($A_0 = \frac{1}{2}a'_0$)	b'_k	A'_k
0	0·037	—	0·019
1	0·012	− 0·004	0·013
2	0·013	− 0·008	0·015
3	0·014	− 0·013	0·019
4	0·016	− 0·020	0·026
5	0·020	− 0·030	0·036
6	0·028	− 0·052	0·058
7	0·056	− 0·121	0·133
8	− 0·371	0·910	0·982
9	− 0·039	0·106	0·113
10	− 0·019	0·059	0·061
11	− 0·012	0·042	0·044
12	− 0·009	0·033	0·034

Note.—It is evident from equations (3) that the results obtained depend upon the phase-relation between the non-harmonic component and the "cycle" chosen. In fact, if analyses are made of successive "cycles" until a complete true cycle is covered, the average of the results would be a true analysis.

2. False cycle.

If, by error, a distance is wrongly estimated to include a cycle of the waveform, some or all of the components may be non-harmonic, and the resulting errors can be computed as in the previous section. It is not considered desirable to give a numerical example of the effect, but a general formula is stated for convenience of reference.

If the false " cycle " chosen has a wavelength q times the true wavelength, the Fourier series becomes :

$$y + A_0 + \sum_k (a_k \cos kqx + b_k \sin kqx) \qquad . \quad . \quad (7)$$

(q will normally be non-integral). The coefficients A'_0, a'_k and b'_k which will be obtained by analysis over the false " cycle "are given by :

$$\left.\begin{aligned} A'_0 &= A_0 + \sum_k \frac{1}{2kq\pi} [a_k \sin 2kq\pi + b_k (1 - \cos 2kq\pi)] \\ a'_k &= \sum_k \frac{q}{\pi k(q^2 - 1)} [a_k \sin 2kq\pi + b_k (1 - \cos 2kq\pi)] \\ b'_k &= \sum_k \frac{1}{\pi k(q^2 - 1)} [- a_k(1 - \cos 2kq\pi) + b_k \sin 2kq\pi] \end{aligned}\right\} \quad . \quad (8)$$

APPENDIX IV

DEMONSTRATION OF FOURIER'S THEOREM

THE rigorous proof of Fourier's Theorem is long and complex; in Chapter VI the normal course was followed—viz. assuming the truth of the theorem and showing how to evaluate the coefficients in the Fourier series. It is, however, possible to demonstrate the theorem in a more direct manner; this demonstration presents several points of interest, and appears to be rigorous except for the assumption "what is true up to the limit is true in the limit."

The following outline is abridged from *Acoustics*, by W. F. Donkin (Clarendon Press, Oxford), by permission of the publishers.

The equation

$$y = \frac{1 - c^2}{1 - 2c\cos(x - a) + c^2}, \quad . \quad . \quad . \quad (1)$$

represents a periodic curve of period 2π radians in x. In this discussion c will be supposed to be restricted to positive values not greater than unity. y then has a maximum value

$$\frac{1 + c}{1 - c},$$

corresponding to the values $x = a \pm 2r\pi$, where r is any integer; these values of x will be termed "critical values." y also has a minimum value

$$\frac{1 - c}{1 + c},$$

corresponding to the values $x = a \pm (2r + 1)\pi$. One cycle of the curve is represented at (a) in the diagram, for the particular case $c = 3/5$. If c is increased towards unity as a limit, the maximum and minimum ordinates tend towards infinite and zero limiting values, as illustrated at (b); and the ordinate corresponding to any value of x other than a critical value tends to zero, so that in the limiting case the curve is the locus of a point which moves along the x-axis but does not pass through the points $[a \pm 2r\pi, 0]$, avoiding these points by traversing the entire positive half of the ordinates erected at these points.

The area included under a cycle of the curve is 2π. This may be established by direct integration of (1); but the result is more easily obtained after first developing y as a power series in c,

$$y = 1 + 2[c\cos(x - a) + c^2\cos 2(x - a) + \ldots], \quad . \quad (2)$$

since

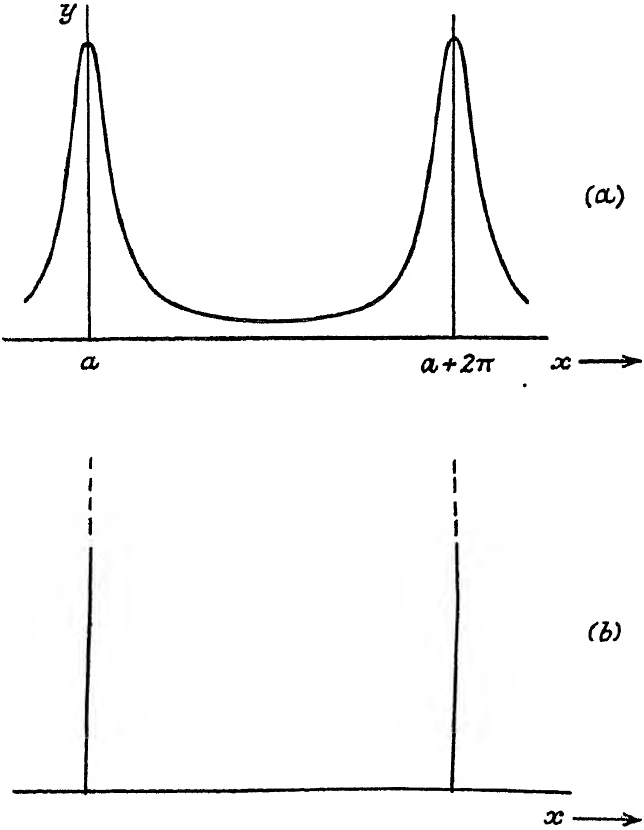
$$\int_a^{x+2\pi} \cos k(x - a)\, dx = 0,$$

if k is any integer 1, 2, 3, etc.

The following inferences are established in the work referred to:

If x_0, x_1 be two values of x, including between them only one critical value (say a), with which neither of them coincides, and if ϵ, ϵ' be any positive constants such that $(a + \epsilon)$ and $(a - \epsilon')$ are both included between x_0 and x_1, then the integrals

$$\int_{x_0}^{a-\epsilon'} y \,.\, dx, \quad \int_{a-\epsilon'}^{a} y \,.\, dx, \quad \int_{a}^{a+\epsilon} y \,.\, dx, \quad \int_{a+\epsilon}^{x_1} y \,.\, dx,$$



Graphs of function (1): (a) $c = 3/5$; (b) $c = 1$.

tend, respectively, towards the values 0, π, π, 0 as c approaches unity, however small ϵ and ϵ' may be. The sum of the four integrals, and the sum of the middle two, both approach 2π as c approaches unity. and the sum of the four integrals is 2π for all values of c if $x_1 - x_0 = 2\pi$,

If now x_0, x_1 are supposed to have any values such that $x_1 - x_0$ is positive but not greater than 2π, and $f(x)$ is any function finite for all values of x from x_0 to x_1 inclusive, then the integral

$$\int_{x_0}^{x_1} f(x)\, y \,.\, dx \quad . \quad . \quad . \quad . \quad . \quad (3)$$

is equal to the product of the integral $\int_{x_0}^{x_1} y \,.\, dx$ and some quantity intermediate between the algebraically least and greatest values assumed by $f(x)$ in this interval. It can then be shown that

(i) If there is no critical value between x_0 and x_1 or at either of these values, the limiting value of (3) as c approaches unity is zero;

(ii) If there is a critical value of x (say $x = a$) between x_0 and x_1, but not coinciding with either of these values, the limiting value of (3) is $2\pi \,.\, f(a)$;

(iii) If both values x_0 and x_1 are critical values, say a and $a + 2\pi$, the limiting value of (3) is

$$\pi[f(a) + f(a + 2\pi)];$$

but if only one of the values x_0 and x_1 is a critical value a, the result is $\pi \,.\, f(a)$.

The above conclusions still hold good if the function $f(x)$ has finite discontinuities at any values of x other than a critical value; if, however, there is such a discontinuity at a critical value $x = a$, the result is slightly modified:

(iv) If a finite discontinuity in $f(x)$ occurs at a critical value a of x, the limiting value of (3) as c approaches unity is

$$\pi[f(x - 0) + f(x + 0)]$$

(see p. 141 for explanation of this notation).

Disregarding for present purposes the exceptional case (iv), it is seen that the limiting value of the integral

$$\int_{x_0}^{x_1} f(x) \,.\, \frac{(1 - c^2)dx}{1 - 2c\cos(x - a) + c^2}, \qquad . \quad . \quad . \quad (4)$$

as c approaches unity is $2\pi \,.\, f(a)$. Interchanging (for convenience) a and x, expanding the fraction in (4) in the series form (2), and assuming that when c equals unity the values of the various integrals equal their limits, the result is obtained:

$$2\pi \,.\, f(x) = \sum_{k=-\infty}^{k=\infty} \int_{a_0}^{a_1} f(a) \cos k(x - a) \,.\, da, \qquad . \quad . \quad (5)$$

where $a_1 - a_0 = 2\pi$. By a simple transformation, equation (5) can be put in the form

$$f(x) = \frac{1}{2\pi}\int_0^{2\pi} f(a) \,.\, da + \frac{1}{\pi}\sum_{k=1}^{k=\infty} \cos kx \int_0^{2\pi} f(a) \,.\, \cos ka \,.\, da$$
$$+ \frac{1}{\pi}\sum_{k=1}^{k=\infty} \sin kx \int_0^{2\pi} f(a) \,.\, \sin ka \,.\, da \qquad (6)$$

which is Fourier's Theorem and includes the Cauchy integral formulæ. If there is a finite discontinuity in $f(x)$ at a particular value of x, then $f(x)$ in (5) and (6) should be replaced by

$$\tfrac{1}{2}[f(x - 0) + f(x + 0)].$$

The reader is referred to Donkin's work for the detailed statement of this demonstration, of which the above is merely an outline.

APPENDIX V

USEFUL TRIGONOMETRICAL FORMULÆ, AND TABLES FOR SYNTHESIS

Some trigonometrical formulæ which occur in the general study of sine-waves are collected below for ease of reference.

1. Compound angles.

$$\sin (A + B) = \sin A \cos B + \cos A \sin B.$$
$$\sin (A - B) = \sin A \cos B - \cos A \sin B.$$
$$\cos (A + B) = \cos A \cos B - \sin A \sin B.$$
$$\cos (A - B) = \cos A \cos B + \sin A \sin B.$$

(multiple angles)

$$\sin 2\theta = 2 \sin \theta \cos \theta.$$
$$\cos 2\theta = 2 \cos^2 \theta - 1 = 1 - 2 \sin^2 \theta = \cos^2 \theta - \sin^2 \theta.$$
$$\sin 3\theta = 3 \sin \theta - 4 \sin^3 \theta.$$
$$\cos 3\theta = 4 \cos^3 \theta - 3 \cos \theta.$$

2. Sum, difference and product formulæ.

$$\sin A + \sin B = 2 \sin \tfrac{1}{2}(A + B) \cos \tfrac{1}{2}(A - B).$$
$$\sin A - \sin B = 2 \cos \tfrac{1}{2}(A + B) \sin \tfrac{1}{2}(A - B).$$
$$\cos A + \cos B = 2 \cos \tfrac{1}{2}(A + B) \cos \tfrac{1}{2}(A - B).$$
$$\cos A - \cos B = 2 \sin \tfrac{1}{2}(A + B) \sin \tfrac{1}{2}(B - A).$$
$$2 \sin A \cos B = \sin (A + B) + \sin (A - B).$$
$$2 \cos A \sin B = \sin (A + B) - \sin (A - B).$$
$$2 \cos A \cos B = \cos (A + B) + \cos (A - B).$$
$$2 \sin A \sin B = \cos (A - B) - \cos (A + B).$$

3. Exponential and series formulæ.

$$2i \sin \theta = e^{i\theta} - e^{-i\theta}.$$
$$2 \cos \theta = e^{i\theta} + e^{-i\theta}.$$

$$\sin \theta = \theta - \frac{\theta^3}{\underline{|3}} + \frac{\theta^5}{\underline{|5}} - \frac{\theta^7}{\underline{|7}} + \text{etc.}$$

$$\cos \theta = 1 - \frac{\theta^2}{\underline{|2}} + \frac{\theta^4}{\underline{|4}} - \frac{\theta^6}{\underline{|6}} + \text{etc.}$$

where θ is in radians and $i^2 = -1$.

4. Relation between radian and degree measurements.

$$1 \text{ radian} = 57{\cdot}2958° = 57° \ 17' \ 45''.$$
$$1° = 0{\cdot}0174533 \text{ radians}.$$
$$1' = 0{\cdot}0002909 \text{ radians}.$$
$$180° = 3{\cdot}14159 \text{ radians}.$$
$$90° = 1{\cdot}570796 \text{ radians}.$$

5. De Moivre's Theorem.

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta,$$

for any value of n, where $i^2 = -1$.

Application.—To express $\cos n\theta$ or $\sin n\theta$ in terms of $\sin\theta$ and $\cos\theta$, n being integral, expand the nth power of $(\cos\theta + i\sin\theta)$ by the binomial theorem, replacing i^2 by -1 wherever it occurs (thus $i^3 = -i$, etc.). Collect the terms which do not include an i; the sum of these terms is $\cos n\theta$. Similarly, $i\sin n\theta$ is the sum of the terms which do include an i, and $\sin n\theta$ is found by dividing this sum by i. Thus:

$$\begin{aligned}&(\cos\theta + i\sin\theta)^3\\ &= \cos^3\theta + 3i\cos^2\theta\sin\theta + 3i^2\cos\theta\sin^2\theta + i^3\sin^3\theta\\ &= \cos^3\theta - 3\cos\theta\sin^2\theta + i(3\cos^2\theta\sin\theta - \sin^3\theta)\\ &= 4\cos^3\theta - 3\cos\theta + i(3\sin\theta - 4\sin^3\theta).\end{aligned}$$

Hence
$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta,$$
$$\sin 3\theta = 3\sin\theta - 4\sin^3\theta.$$

6. Small angles.

θ and ϕ being small and measured in radians, the following relations are *approximately* true:

$$\begin{aligned}\sin\theta &= \theta,\\ \cos\theta &= 1 - \tfrac{1}{2}\theta^2 \text{ or, less accurately, } = 1.\\ \sin(\theta+\phi) &= \theta + \phi,\\ \sin(\theta-\phi) &= \theta - \phi,\\ \cos(\theta+\phi) &= 1 - \theta\phi - \tfrac{1}{2}(\theta^2+\phi^2),\\ \cos(\theta-\phi) &= 1 + \theta\phi - \tfrac{1}{2}(\theta^2+\phi^2).\end{aligned}$$

TABLE I

$\theta°$.	Sin θ.	Cos θ
0	0	1
15	0·2588	0·9659
30	0·5000	0·8660
45	0·7071	0·7071
60	0·8660	0·5000
75	0·9659	0·2588
90	1	0
105	0·9659	− 0·2588
120	0·8660	− 0·5000
135	0·7071	− 0·7071
150	0·5000	− 0·8660
165	0·2588	− 0·9659
180	0	− 1
195	− 0·2588	− 0·9659
210	− 0·5000	− 0·8660
etc.	etc.	etc.

TABLE I

Values of sine and cosine harmonic terms up to the 16th at intervals of $3{\cdot}75° = 360°/96$

[Harmonic number k, angle $3{\cdot}75\ r°$; table gives $\sin k\ (3{\cdot}75\ r°)$ and $\cos k\ (3{\cdot}75\ r°)$, S ≡ sine, C ≡ cosine]

r =	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
k																	
1 S	·065	·131	·195	·259	·321	·383	·442	·500	·556	·609	·659	·707	·752	·793	·832	·866	S 1
1 C	·998	·991	·981	·966	·947	·924	·897	·866	·832	·793	·752	·707	·659	·609	·556	·500	C 1
2 S	·131	·259	·383	·500	·609	·707	·793	·866	·924	·966	·991	1	·991	·966	·924	·866	S 2
2 C	·991	·966	·924	·866	·793	·707	·609	·500	·383	·259	·131	0	−·131	−·259	−·383	−·500	C 2
3 S	·195	·383	·556	·707	·832	·924	·981	1	·981	·924	·832	·707	·556	·383	·195	0	S 3
3 C	·981	·924	·832	·707	·556	·383	·195	0	−·195	−·383	−·556	−·707	−·832	−·924	−·981	−1	C 3
4 S	·259	·500	·707	·866	·966	1	·966	·866	·707	·500	·259	0	−·259	−·500	−·707	−·866	S 4
4 C	·966	·866	·707	·500	·259	0	−·259	−·500	−·707	−·866	−·966	−1	−·966	−·866	−·707	−·500	C 4
5 S	·321	·609	·832	·966	·998	·924	·752	·500	·195	−·131	−·442	−·707	−·897	−·991	−·981	−·866	S 5
5 C	·947	·793	·556	·259	−·065	−·383	−·659	−·866	−·981	−·991	−·897	−·707	−·442	−·131	·195	·500	C 5
6 S	·383	·707	·924	1	·924	·707	·383	0	−·383	−·707	−·924	−1	−·924	−·707	−·383	0	S 6
6 C	·924	·707	·383	0	−·383	−·707	−·924	−1	−·924	−·707	−·383	0	·383	·707	·924	1	C 6
7 S	·442	·793	·981	·966	·752	·383	−·065	−·500	−·832	−·991	−·947	−·707	−·321	·131	·556	·866	S 7
7 C	·897	·609	·195	−·259	−·659	−·924	−·998	−·866	−·556	−·131	·321	·707	·947	·991	·832	·500	C 7
8 S	·500	·866	1	·866	·500	0	−·500	−·866	−1	−·866	−·500	0	·500	·866	1	·866	S 8
8 C	·866	·500	0	−·500	−·866	−1	−·866	−·500	0	·500	·866	1	·866	·500	0	−·500	C 8
9 S	·556	·924	·981	·707	·195	−·383	−·832	−1	−·832	−·383	·195	·707	·981	·924	·556	0	S 9
9 C	·832	·383	−·195	−·707	−·981	−·924	−·556	0	·556	·924	·981	·707	·195	−·383	−·832	−1	C 9
10 S	·609	·966	·924	·500	−·131	−·707	−·991	−·866	−·383	·259	·793	1	·793	·259	−·383	−·866	S 10
10 C	·793	·259	−·383	−·866	−·991	−·707	−·131	·500	·924	·966	·609	0	−·609	−·966	−·924	−·500	C 10
11 S	·659	·991	·832	·259	−·442	−·924	−·947	−·500	·195	·793	·998	·707	·065	−·609	−·981	−·866	S 11
11 C	·752	·131	−·556	−·966	−·897	−·383	·321	·866	·981	·609	−·065	−·707	−·998	−·793	−·195	·500	C 11
12 S	·707	1	·707	0	−·707	−1	−·707	0	·707	1	·707	0	−·707	−1	−·707	0	S 12
12 C	·707	0	−·707	−1	−·707	0	·707	1	·707	0	−·707	−1	−·707	0	·707	1	C 12
13 S	·752	·991	·556	−·259	−·897	−·924	−·321	·500	·981	·793	·065	−·707	−·998	−·609	·195	·866	S 13
13 C	·659	−·131	−·832	−·966	−·442	·383	·947	·866	·195	−·609	−·998	−·707	·065	·793	·981	·500	C 13
14 S	·793	·966	·383	−·500	−·991	−·707	·131	·866	·924	·259	−·609	−1	−·609	·259	·924	·866	S 14
14 C	·609	−·259	−·924	−·866	−·131	·707	·991	·500	−·383	−·966	−·793	0	·793	·966	·383	−·500	C 14
15 S	·832	·924	·195	−·707	−·981	−·383	·556	1	·556	−·383	−·981	−·707	·195	·924	·832	0	S 15
15 C	·556	−·383	−·981	−·707	·195	·924	·832	0	−·832	−·924	−·195	·707	·981	·383	−·556	−1	C 15
16 S	·866	·866	0	−·866	−·866	0	·866	·866	0	−·866	−·866	0	·866	·866	0	−·866	S 16
16 C	·500	−·500	−1	−·500	·500	1	·500	−·500	−1	−·500	·500	1	·500	−·500	−1	−·500	C 16
r =	95	94	93	92	91	90	89	88	87	86	85	84	83	82	81	80	

Note: for values of r from 48 to 96, change sign of every sine (S) term.

TABLE II—(*continued*)

Values of sine and cosine harmonic terms up to the 16th at intervals of $3{\cdot}75° = 360°/96$

[Harmonic number k, angle $3{\cdot}75\ r°$; table gives $\sin k\ (3{\cdot}75\ r°)$ and $\cos k\ (3{\cdot}75\ r°)$, S ≡ sine, C ≡ cosine]

$r =$		17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32		
k																			
1	S	·897	·924	·947	·966	·981	·991	·998	1	·998	·991	·981	·966	·947	·924	·897	·866	S	1
	C	·442	·383	·321	·259	·195	·131	·065	0	− ·065	− ·131	− ·195	− ·259	− ·321	− ·383	− ·442	− ·500	C	
2	S	·793	·707	·609	·500	·383	·259	·131	0	− ·131	− ·259	− ·383	− ·500	− ·609	− ·707	− ·793	− ·866	S	2
	C	− ·609	− ·707	− ·793	− ·866	− ·924	− ·966	− ·991	− 1	− ·991	− ·966	− ·924	− ·866	− ·793	− ·707	− ·609	− ·500	C	
3	S	− ·195	− ·383	− ·556	− ·707	− ·832	− ·924	− ·981	− 1	− ·981	− ·924	− ·832	− ·707	− ·556	− ·383	− ·195	0	S	3
	C	− ·981	− ·924	− ·832	− ·707	− ·556	− ·383	− ·195	0	·195	·383	·556	·707	·832	·924	·981	1	C	
4	S	− ·966	− 1	− ·966	− ·866	− ·707	− ·500	− ·259	0	·259	·500	·707	·866	·966	1	·966	·866	S	4
	C	− ·259	0	·259	·500	·707	·866	·966	1	·966	·866	·707	·500	·259	0	− ·259	− ·500	C	
5	S	− ·659	− ·383	− ·065	·259	·556	·793	·947	1	·947	·793	·556	·259	− ·065	− ·383	− ·659	− ·866	S	5
	C	·752	·924	·998	·966	·832	·609	·321	0	− ·321	− ·609	− ·832	− ·966	− ·998	− ·924	− ·752	− ·500	C	
6	S	·383	·707	·924	1	·924	·707	·383	0	− ·383	− ·707	− ·924	− 1	− ·924	− ·707	− ·383	0	S	6
	C	·924	·707	·383	0	− ·383	− ·707	− ·924	− 1	− ·924	− ·707	− ·383	0	·383	·707	·924	1	C	
7	S	·998	·924	·659	·259	− ·195	− ·609	− ·897	− 1	− ·897	− ·609	− ·195	·259	·659	·924	·998	·866	S	7
	C	·065	− ·383	− ·752	− ·966	− ·981	− ·793	− ·442	0	·442	·793	·981	·966	·752	·383	− ·065	− ·500	C	
8	S	·500	0	− ·500	− ·866	− 1	− ·866	− ·500	0	·500	·866	1	·866	·500	0	− ·500	− ·866	S	8
	C	− ·866	− 1	− ·866	− ·500	0	·500	·866	1	·866	·500	0	− ·500	− ·866	− 1	− ·866	− ·500	C	
9	S	− ·556	− ·924	− ·981	− ·707	− ·195	·383	·832	1	·832	·383	− ·195	− ·707	− ·981	− ·924	− ·556	0	S	9
	C	− ·832	− ·383	·195	·707	·981	·924	·556	0	− ·556	− ·924	− ·981	− ·707	− ·195	·383	·832	1	C	
10	S	− ·991	− ·707	− ·131	·500	·924	·966	·609	0	− ·609	− ·966	− ·924	− ·500	·131	·707	·991	·866	S	10
	C	·131	·707	·991	·866	·383	− ·259	− ·793	− 1	− ·793	− ·259	·383	·866	·991	·707	·131	− ·500	C	
11	S	− ·321	·383	·897	·966	·556	− ·131	− ·752	− 1	− ·752	− ·131	·556	·966	·897	·383	− ·321	− ·866	S	11
	C	·947	·924	·442	− ·259	− ·832	− ·991	− ·659	0	·659	·991	·832	·259	− ·442	− ·924	− ·947	− ·500	C	
12	S	·707	1	·707	0	− ·707	− 1	− ·707	0	·707	1	·707	0	− ·707	− 1	− ·707	0	S	12
	C	·707	0	− ·707	− 1	− ·707	0	·707	1	·707	0	− ·707	− 1	− ·707	0	·707	1	C	
13	S	·947	·383	− ·442	− ·966	− ·832	− ·131	·659	1	·659	− ·131	− ·832	− ·966	− ·442	·383	·947	·866	S	13
	C	− ·321	− ·924	− ·897	− ·259	·556	·991	·752	0	− ·752	− ·991	− ·556	·259	·897	·924	·321	− ·500	C	
14	S	·131	− ·707	− ·991	− ·500	·383	·966	·793	0	− ·793	− ·966	− ·383	·500	·991	·707	− ·131	− ·866	S	14
	C	− ·991	− ·707	·131	·866	·924	·259	− ·609	− 1	− ·609	·259	·924	·866	·131	− ·707	− ·991	− ·500	C	
15	S	− ·832	− ·924	− ·195	·707	·981	·383	− ·556	− 1	− ·556	·383	·981	·707	− ·195	− ·924	− ·832	0	S	15
	C	− ·556	·383	·981	·707	− ·195	− ·924	− ·832	0	·832	·924	·195	− ·707	− ·981	− ·383	·556	1	C	
16	S	− ·866	0	·866	·866	0	− ·866	− ·866	0	·866	·866	0	− ·866	− ·866	0	·866	·866	S	16
	C	·500	1	·500	− ·500	− 1	− ·500	·500	1	·500	− ·500	− 1	− ·500	·500	1	·500	− ·500	C	
$r =$		79	78	77	76	75	74	73	72	71	70	69	68	67	66	65	64		

Note.—For values of r from 48 to 96, change sign of every Sine (S) term.

TABLE II—(*continued*)

Values of sine and cosine harmonic terms up to the 16th at intervals of 3·75° = 360°/96

[Harmonic number k, angle 3·75 r° ; table gives sin k (3·75 r°) and cos k (3·75 r°), S ≡ sine, C ≡ cosine]

r =	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	
k																	
1 S	·832	·793	·752	·707	·659	·609	·556	·500	·442	·383	·321	·259	·195	·131	·065	0	S 1
C	− ·556	− ·609	− ·659	− ·707	− ·752	− ·793	− ·832	− ·866	− ·897	− ·924	− ·947	− ·966	− ·981	− ·991	− ·998	− 1	C
2 S	− ·924	− ·966	− ·991	− 1	− ·991	− ·966	− ·924	− ·866	− ·793	− ·707	− ·609	− ·500	− ·383	− ·259	− ·131	0	S 2
C	− ·383	− ·259	− ·131	0	·131	·259	·383	·500	·609	·707	·793	·866	·924	·966	·991	1	C
3 S	·195	·383	·556	·707	·832	·924	·981	1	·981	·924	·832	·707	·556	·383	·195	0	S 3
C	·981	·924	·832	·707	·556	·383	·195	0	− ·195	− ·383	− ·556	− ·707	− ·832	− ·924	− ·981	− 1	C
4 S	·707	·500	·259	0	− ·259	− ·500	− ·707	− ·866	− ·966	− 1	− ·966	− ·866	− ·707	− ·500	− ·259	0	S 4
C	− ·707	− ·866	− ·966	− 1	− ·966	− ·866	− ·707	− ·500	− ·259	0	·259	·500	·707	·866	·966	1	C
5 S	− ·981	− ·991	− ·897	− ·707	− ·442	− ·131	·195	·500	·752	·924	·998	·966	·832	·609	·321	0	S 5
C	− ·195	·131	·442	·707	·897	·991	·981	·866	·659	·383	·065	− ·259	− ·556	− ·793	− ·947	− 1	C
6 S	·383	·707	·924	1	·924	·707	·383	0	− ·383	− ·707	− ·924	− 1 .	− ·924	− ·707	− ·383	0	S 6
C	·924	·707	·383	0	− ·383	− ·707	− ·924	− 1	− ·924	− ·707	− ·383	0	·383	·707	·924	1	C
7 S	·556	·131	− ·321	− ·707	− ·947	− ·991	− ·832	− ·500	− ·065	·383	·752	·966	·981	·793	·442	0	S 7
C	− ·832	− ·991	− ·947	− ·707	− ·321	·131	·556	·866	·998	·924	·659	·259	− ·195	− ·609	− ·897	− 1	C
8 S	− 1	− ·866	− ·500	0	·500	·866	1	·866	·500	0	− ·500	− ·866	− 1	− ·866	− ·500	0	S 8
C	0	·500	·866	1	·866	·500	0	− ·500	− ·866	− 1	− ·866	− ·500	0	·500	·866	1	C
9 S	·556	·924	·981	·707	·195	− ·383	− ·832	− 1	− ·832	− ·383	·195	·707	·981	·924	·556	0	S 9
C	·832	·383	− ·195	− ·707	− ·981	− ·924	− ·556	0	·556	·924	·981	·707	·195	− ·383	− ·832	− 1	C
10 S	·383	− ·259	− ·793	− 1	− ·793	− ·259	·383	·866	·991	·707	·131	− ·500	− ·924	− ·966	− ·609	0	S 10
C	− ·924	− ·966	− ·609	0	·609	·966	·924	·500	− ·131	− ·707	− ·991	− ·866	− ·383	·259	·793	1	C
11 S	− ·981	− ·609	·065	·707	·998	·793	·195	− ·500	− ·947	− ·924	− ·442	·259	·832	·991	·659	0	S 11
C	·195	·793	·998	·707	·065	− ·609	− ·981	− ·866	− ·321	·383	·897	·966	·556	− ·131	− ·752	− 1	C
12 S	·707	1	·707	0	− ·707	− 1	− ·707	0	·707	1	·707	0	− ·707	− 1	− ·707	0	S 12
C	·707	0	− ·707	− 1	− ·707	0	·707	1	·707	0	− ·707	− 1	− ·707	0	·707	1	C
13 S	·195	− ·609	− ·998	− ·707	·065	·793	·981	·500	− ·321	− ·924	− ·897	− ·259	·556	·991	·752	0	S 13
C	− ·981	− ·793	− ·065	·707	·998	·609	− ·195	− ·866	− ·947	− ·383	·442	·966	·832	·131	− ·659	− 1	C
14 S	− ·924	− ·259	·609	1	·609	− ·259	− ·924	− ·866	− ·131	·707	·991	·500	− ·383	− ·966	− ·793	0	S 14
C	·383	·966	·793	0	− ·793	− ·966	− ·383	·500	·991	·707	− ·131	− ·866	− ·924	− ·259	·609	1	C
15 S	·832	·924	·195	− ·707	− ·981	− ·383	·556	1	·556	− ·383	− ·981	− ·707	·195	·924	·832	0	S 15
C	·556	− ·383	− ·981	− ·707	·195	·924	·832	0	− ·832	− ·924	− ·195	·707	·981	·383	− ·556	− 1	C
16 S	0	− ·866	− ·866	0	·866	·866	0	− ·866	− ·866	0	·866	·866	0	− ·866	− ·866	0	S 16
C	− 1	− ·500	·500	1	·500	− ·500	− 1	− ·500	·500	1	·500	− ·500	− 1	− ·500	·500	1	C
r =	63	62	61	60	59	58	57	56	55	54	53	52	51	50	49	48	

Note.—For values of r from 48 to 96, change sign of every sine (S) term.

7. Differentials and integrals.

In the following formulæ, n may have any value :

$$\frac{d(\sin nx)}{dx} = n \,.\, \cos nx$$

$$\frac{d(\cos nx)}{dx} = -\, n \,.\, \sin nx$$

$$\int \sin nx \,.\, dx = \frac{-\cos nx}{n} + \mathrm{C}$$

$$\int \cos nx \,.\, dx = \frac{\sin nx}{n} + \mathrm{C}$$

8. Tables.

Table I gives the values of the sine and cosine functions at intervals of 15°.

Table II gives the values of the first sixteen sine and cosine harmonics, at intervals of $3\frac{3}{4}° = 360°/96$. The actual angle for which the values are tabulated is $3\frac{3}{4} \,.\, kr°$. It must be noted that for values of r between 48 and 96, the signs of the *sine* harmonics are to be reversed.

GLOSSARY

The most frequently recurring descriptive terms and expressions are defined briefly below. The numbers in parentheses refer to the pages on which the terms are first defined or explained in the text.

Alternance (14). A property possessed exclusively by certain *periodic* functions, the Fourier series for which contain only the odd harmonics; an alternant function or waveform is such that each half-cycle is the same as the preceding half-cycle, except for reversal of sign.

Amplitude (9). Half the total excursion of a sine-wave; i.e. the maximum deviation from the mean or centre-line.

Double amplitude (9). The total excursion or overall height of a wave; in a sine-wave, twice the amplitude.

Anti-phase (10). Of two waves of the same frequency, the condition in which the phase-difference is 180°, so that crests of one always occur simultaneously with troughs in the other, and *vice versa;* of two waves of different frequencies, temporary anti-phase is the condition in which a crest of one occurs simultaneously with a trough in the other, or in which zero values of both waves occur simultaneously, the curves diverging one to either side of the centre-line, so that the gradients of the two curves are of opposite sign. Beats of the same frequency in two waves recorded simultaneously are said to be anti-phased when bulges in one coincide with waists in the other, and *vice versa.*

Apparent low-frequency surge (38). The illusory indication of a low-frequency component in a wave whose two true components have a frequency ratio which is approximately integral, or due to beating between two high frequency components superimposed upon a low frequency component.

Basic variable (9). The variable whose value determines that of a function; x in $y=f(x)$. Sometimes called the "independent variable," "argument" or "time-base."

Beating (20). The regular alternate increase and decrease in the amplitude of a wave, caused by the addition of another component of nearly equal frequency.

Beat frequency (21). The number of complete beats, i.e. of swellings, occurring in unit time; equal to the difference between the frequencies of the components.

Bulge (25). That part of a beating waveform where the amplitude is greatest.

Cauchy integrals (144). The formulæ by which the Fourier coefficients may be calculated.

Centre-line (9). The line representing the average value of a periodic variation over a cycle.

Coefficients (143). The two constants which, with the frequency, completely determine a sinusoidal wave; either the amplitude and phase, or the amplitudes of the sine and cosine components.

Components (15). The sine-waves from which a complex variation is built up, and into which it can be decomposed. Also, of a single sinusoidal wave, the constituent parts which are in-phase with a sine and a cosine wave of the same frequency, referred to the same datum for the basic variable.

Constant term (49). The average value of a periodic variation, taken over a cycle.

Cosine component (14). A sinusoidal wave which has zero values at the end of the first and third quarter-cycles.

Crest (9). That part of a waveform which corresponds to a maximum in the variation.

Cycle (9). The complete pattern of a periodic variation, which is repeated at regular intervals equal to the length of the cycle.

Sub-cycle (120). A part of a complete cycle; a waveform is divided into equal sub-cycles in the method of superposition.

Datum (49). A convenient line or level from which the value of the basic or the dependent variable is measured.

Dependent variable (138). The variable whose value is given in terms of the basic variable: y in $y=f(x)$. The quantity whose periodic variation generates a periodic wave or function.

Disguised alternance or skew-symmetry (58). The property exhibited by a wave or variation consisting of a constant term added to an alternant or skew-symmetrical wave or variation.

Envelope (26). An auxiliary line constructed on a waveform, for the purpose of analysis by the inspection method: it passes through all the crests, or all the troughs, or touches the wave near all the crests or all the troughs.

Envelope mean (40). The line midway between the crest and trough envelopes.

Envelope strip (27). The area between the crest and trough envelopes,

Even components (51). Harmonics whose reference numbers are even; hitherto used as a synonym for "symmetrical," but not so in this book.

Excursion (78). The variation of the dependent variable from the mean value, or of a waveform from the centre-line.

Existence function (154). A function having unit value for values of the basic variable within certain ranges, and zero value for all other values of the basic variable.

Film speed (45). The rate at which the photographic film strip moves through a recording camera.

Fourier series (2). A series of sinusoidal components, whose frequencies form the ratios .1 : 2 : 3 : etc., representing a periodic function. Strictly, a series with no limit to the possible number of components, but often used for limited series.

Fourier's Theorem (143). The basis of all waveform analysis; states that any periodic variation fulfilling certain conditions of continuity can be expressed in the form of a Fourier series (*q.v.*).

Frequency (12). The reciprocal of the period of a periodic variation, or the number of cycles occurring in unit time.

Fundamental (48). Referring to the cycle of a complex variation. The fundamental frequency is the frequency of the variation, and the fundamental component is the harmonic with this frequency.

Gibbs' phenomenon (177). The deviation of the sum of a Fourier series (from expansion of a periodic variation) from the value of the variation itself, where this is discontinuous.

Half-range series (151). A Fourier series containing only sine components, or only cosine components.

Harmonic (48). Any component of a Fourier series, but usually referring to any component other than the fundamental; the frequency of a harmonic is an exact multiple of the fundamental frequency.

Harmonic analysis. The process of decomposing a complex variation into its constituent sine waves.

Harmonic reference number (48). The number expressing the ratio of the frequency of the harmonic to the fundamental frequency, or the number of cycles of the harmonic included in a cycle of the complex variation.

In-phase (11). Of two waves of the same frequency, the condition in which the phase-difference is zero; of two waves of different frequencies, temporary in-phase is the condition in which crests or troughs in both waves occur simultaneously, or in which zero values of both waves occur simultaneously, the two curves lying to the same side of the centre-line in the neighbourhood of the point in question, so that their gradients are of the same sign. Beats of the same frequency in two waves recorded simultaneously are said to be in-phase when bulges in one coincide with bulges in the other.

Inspection method (83). The method of waveform analysis wherein the amplitudes, frequencies and phase-angles of the predominant components are determined by consideration of the envelopes and other general properties of the waveform.

Limited series (184). An approximate representation of a periodic variation by a Fourier series with a limited number of components; in general, such a series only represents the variation at a finite number of points in the cycle.

Magnification (49). The ratio of the excursion of a recorded waveform to the excursion of the physical variation represented thereby.

Major component (22). That one of two components which has the larger amplitude (see note on p. 22).

Minor component (25). That one of two components which has the smaller amplitude (see note on p. 22).

Non-harmonic component (3). An extraneous component whose frequency is not an exact multiple of the fundamental frequency of a given variation, and which does not therefore belong to the harmonic series for that variation.

Odd components (53). Harmonics whose reference numbers are odd; hitherto used as a synonym for "skew-symmetrical," but not so in this book.

Order number (68). The ratio of the frequency of a wave to some standard frequency; often used in connection with vibrations in engines and propellers, the standard frequency being the rotational speed of the engine or propeller: the third engine order has three cycles per engine revolution.

Ordinates (184). The lengths of lines drawn from a datum to a waveform, representing the differences between values of the variation at various parts of the cycle and the datum value.

Paper speed (45). The speed at which the paper strip passes through a recording instrument.

Peak (9). A crest or a trough, i.e. a turning-point in the variation.

Period (9). Strictly, the *time*-interval occupied by a cycle. Also used as a synonym for " wavelength."

Periodic function or variation (1). A function or variation which repeats after successive equal intervals of time, or with successive equal increments of the basic variable.

Phase-angle (9). A measurement relating the position of a sinusoidal wave to an arbitrary datum of the basic variable.

Phase-difference (10). The difference between the phase-angles of two waves of the same frequency.

Phase-velocity (12). The angular velocity of the rotating vector which generates a sine-wave; the ratio of the velocity amplitude to the displacement amplitude, or similarly, of acceleration to velocity, for a sinusoidal displacement.

Reference number (48). The number expressing the ratio of the frequency of a harmonic to the frequency of the fundamental component, i.e. to the frequency of the periodic variation.

Riemann's Theorem (183). The proposition that the (unlimited) Fourier series expansion of a periodic variation is unique.

Sine-wave (8). The wave generated by the uniform rotation of a vector; strictly, that species of such a wave which is zero at the beginning of a cycle, as opposed to a cosine component (*q.v.*), but used generally to denote any such wave displaced by any phase-angle.

Sinusoidal. Varying in the manner of a sine-wave.

Skew-symmetry (13). A function is skew-symmetrical about a value X of the basic variable x if its values for any pair of values of x equidistant from X are equal in magnitude but opposite in sign, so that if a graph of the function were rotated through 180° about the point $x = X$ the appearance would be unaltered.

Superposition (120). That method of analysis wherein various sets of components are separated by the superposition of sub-cycles of the variation, corresponding values in the various sub-cycles being added together.

Symmetry (13). A function is symmetrical about a value X of the basic variable x if its values for any pair of values of x equidistant from X are equal, so that the part of the graph to one side of the line $x = X$ is the mirror-image in this line of the part to the other side.

Synthesis (4). The process of forming a complex variation by the addition of sinusoidal components.

Time-base (45). Sometimes used as a synonym for " basic variable," but strictly refers only to cases where the basic variable is time.

Trough (9). That part of a waveform corresponding to a minimum in the variation.

Waist (25). That part of a beating waveform where the amplitude is least.

Waveform. A graph representing a periodic variation or the sum of a number of periodic variations.

Wavelength (14). The length of a cycle of a periodic waveform.

ANSWERS TO EXERCISES

Chapter I (page 43).

(1) 3·353, 17°. (2) 3·635, 4°. (3) 5·000, 53°. (4) 3·500, 30°. (5) 182·2, 65°.

(6)–(10). (Assuming t to be in hundredths of a second).

	Beat Frequency (C.P.M.)	Envelope amp. (±)	Separation greater at	Beats per cycle	Crests per cycle
(6)	955	5	waist	1	12
(7)	57·3	0·8	waist	3	50
(8)	477	3	bulge	1	17
(9)	1910	9	—	1	4
(10)	191	3·8	bulge	1	28

(11)-(15). (13) and (15) symmetrical about 0 and 180°, (11), (12) and (14) about 90° and 270°; (12) and (14) skew-symmetrical about 0 and 180°, (15) about 90° and 270°; (12), (14) and (15) alternant.

(16)-(20). Assuming t to be in hundredths of a second, frequencies in C.P.M.:—

(16)—(17) 955. (18) 382. (19)—(20) 477.

Chapter II (page 63).

(1) 45 m.; 17·3 mm.; 5200 C.P.M.

(2) 2, 1 and 0·5 ins.; 8th and 24th harmonics.

(3) (*a*) 46·8 ins.; 78th and 48th harmonics. (*b*) 23·4 ins.; 39th and 24th harmonics.

(4) (*a*) and (*b*): 3·6 ins.; 6th and 4th harmonics.

(5) $3{\cdot}2 + 1{\cdot}5 \sin x_1 + 2{\cdot}7 \sin (2x_1 + 45°) + 0{\cdot}8 \sin (4x_1 - 9°)$, where $x_1 = x - 19°$.

(6) $\sin (x_1 + 15°) + 2{\cdot}5 \sin (2x_1 - 45°) + \cos 3x_1 + 3{\cdot}6 \sin (4x_1 + 60°)$, where $x_1 = x - 15°$.

(7) (*a*) 85°; (*b*) 90°; (*c*) 120°; (*d*) 0.

(8) (i), (iii) and (iv) alternant; (iii) and (v) symmetrical about 0 and 180°, (ii) about 90° and 270°; (vi) skew-symmetrical about 0 and 180°, (iii) about 90° and 270°.

BIBLIOGRAPHY

In the list of references, the letters (A) to (E) designate the following general works :

(A) Carslaw, " Fourier's Series and Integrals," Macmillan, 2nd edition.
(B) Eagle, " Fourier's Theorem and Harmonic Analysis," Longman, 1925.
(C) Whittaker and Watson, " A Course in Modern Analysis," Cambridge, 1927.
(D) Karman and Biot, " Mathematical Methods in Engineering," McGraw Hill, 1940.
(E) Manley, " Fundamentals of Vibration Study," Chapman and Hall, 1942.

Introduction :

1. Piaggio, " Differential Equations," Bell, 1925, p. 54.
2. Runge, " Theorie und Praxis der Reihen," Leipzig, 1904, p. 147 ; also Zeits. Math., 1903, p. 443 and 1905, p. 117.

Chapter I :

1. (E) p. 9.
2. (E) p. 25.
3. (E) p. 95.
4. Ker Wilson, " Practical Solution of Torsional Vibration Problems," Chapman and Hall, 1941, Vol. II, p. 326.

Chapter V :

1. Humphrey, " Advanced Mathematics," Oxford, 1929, Part II, p. 104.
2. (E) Appendix I.

Chapter VI :

1. (B) Chapter IX.
2. (A) Historical Introduction.
3. (A) Chapter VII.
4. (C) Chapter IX.
5. (A) Chapter VIII.
6. (B) Chapter VI.
7. (C) Chapter IX.
8. (C) p. 180.
9. (A) Chapter IX.
10. Hussman, " Vibrations in Helical Valve Springs," Jahrbuch der deutschen Luftfahrtforschung, 1938, p. 119.
11. (D) pp. 382, 388 and 405.
12. (A) Chapter X.
13. (D) p. 343.

Chapter VII :

1. Manley, " The Transient State and the Analysis of Shock Excitations," Engineering, Oct. 22 and Nov. 5, 1943.
2. (C) p. 186.

3. Stansfield, "Measurement of Torsional Vibrations," *Proc. I. Mech. E.*, 1942, Vol. CXLVIII, p. 187.
4. Runge—see *Introduction*, ref. 2.
5. "Modern Instruments and Methods of Calculation" (Handbook of the Napier Tercentenary Exhibition), Bell, p. 234.
6. Fischer Hinnen, *Elektrotechnische Zeits.*, 1901, p. 369.
7. Lewis, "A Method of Harmonic Analysis," *Trans. Amer. Soc. Mech. E.*, 1935, p. A 137.

Chapter VIII:

1. Napier handbook (see *Chapter VII*, reference 5), p. 220.
2. Johnson, "A Cathode-ray Method of Wave Analysis," *Elec. Eng.* (Trans.) 1941, p. 1032.
3. Bell Laboratories Record, Vol. XVIII, No. 1, p. 28.
4. Catalogue of the General Radio Corporation of America.
5. Stansfield (see *Chapter VII*, reference 3), p. 175.
6. Timoshenko, "Vibration Problems in Engineering," Constable, 2nd edition, 1937, p. 443.
7. Scientific Computing Service, Ltd., 23 Bedford Square, London, W.C. 1.

Chapter IX:

1. Ker Wilson (see *Chapter I*, reference 4) Vol. II., Chapter 8.
2. Bernhard, "Mechanical Vibrations," Pitman, 1942, Part II.

Chapter X:

1. Osgood, "Mechanics," Macmillan, 1937, plate following p. 190.
2. Harwood and Manley, "The Determination of the Phase Angle between Two Sinusoidal Variations of Equal Frequency by Means of a Lissajou Figure," Math. Gazette, Oct., 1943, Vol. XXVII, No. 276.
3. Parr, "The Cathode-ray Tube," Chapman and Hall, 1944, p. 57.

Appendices:

1. (E) p. 5.
2. (C) p. 224.
3. (E) p. 95.
4. (C) p. 186.

INDEX.

Figures refer to pages, except that where an item is the subject of an entire chapter or section this fact is indicated in brackets : e.g., " Convergence of Fourier Series " is the subject of Chapter VI, section 7, commencing on page 172, and the indication is " (VI, 7) 172 ".

Lists of Examples in text, Formulæ, Practical Notes and Tables are included under the letter headings E, F, P, T.